Are Color Experiences the Same across the Visual Field?

Ariel Zeleznikow-Johnston[1], Yasunori Aizawa[2,3], Makiko Yamada[2], and Naotsugu Tsuchiya[1,4,5]

Abstract

■ It seems obvious to laypeople that neurotypical humans experience color equivalently across their entire visual field. To some neuroscientists, psychologists, and philosophers, though, this claim has been met with skepticism, as neurophysiological evidence indicates the mechanisms that support color perception degrade with eccentricity. However, the argument that this entails altered color experience in peripheral vision is not universally accepted. Here, we address whether color experience is essentially equivalent between central and peripheral vision. To assess this, we will obtain similarity relationships between color experiences across the visual field using both online and laboratory-based far-field displays, while removing the confounds of saccades, memory, and expectation about color experiences. Our experiment was designed to provide clear evidence that would favor either unchanged or altered color experience relationships in the periphery. Our results are consistent with lay people's phenomenological reports: Color experiences, as probed by similarity relationships in central vision and the far field (60°), are equivalent when elicited by large stimuli. These findings challenge the widespread view in philosophy and cognitive science that peripheral color experiences are illusory, and are discussed in the context of their related neurophysiological, psychophysical, and philosophical literature. ■

INTRODUCTION

It seems obvious to laypeople that neurotypical humans experience color across their entire visual field. Any distinct point in the visual field has some color associated with it, without a sense that the possible colors one could experience differs at different locations. To neuroscientists though, this claim is sometimes met with skepticism because of differences in perceptual performance across the visual field. Performance on tasks using small fixed-size stimuli decreases with eccentricity (Strasburger, Rentschler, & Juttner, 2011), and the required color contrast for differences to be detected increases (Hansen, Pracejus, & Gegenfurtner, 2009). There are associated neurophysiological changes with increasing retinal eccentricity, such as declines in the density of color-sensitive photoreceptors, cones (Curcio, Sloan, Kalina, & Hendrickson, 1990). This behavioral and physiological evidence is consistent with the claim that abilities to perceive color differ across the visual field. However, some researchers go even further, believing that color perception is not just different but actually degraded in peripheral vision compared with foveal vision, and that this entails a degraded experience of color in the periphery. Comments such as "perceptual experience lacks a surprising amount of color" (Cohen & Rubenstein, 2020), "introspection [about peripheral color] cannot be truly reliable" (Giron, Lau, & Knotts, 2018), "our intuitive sense of a rich, colorful visual world is largely incorrect" (Cohen, Botch, & Robertson, 2020), and "it seems that our color vision goes right way out to the edge of our vision. It doesn't." (Carroll & Dennett, 2020) are common. These authors believe that the unreliable nature of introspection into perception should lead us to be skeptical of reports about conscious experience of peripheral color. We seek to test the validity of these color-sceptic claims by determining whether color experiences are equivalent across the visual field.

The most dramatic support for the skeptical view comes from a recent study showing that people can be entirely unaware of gradual, yet complete, removal of peripheral visual color inputs under seminaturalistic viewing conditions (Cohen et al., 2020). Using a virtual-reality setup, the authors demonstrated that the majority of people do not notice when color is removed from eccentricities greater than 20 degrees of visual angle (DVA). Further support comes from studies that asked participants to assign a color to a given stimulus presented for short durations (380–1000 msec) at varying eccentricities. Using the same stimuli presented centrally or peripherally, reports of a roughly 50% decrease in perceived saturation at around 40 DVA are observed, as well as shifts in the assigned hue and increased contrast required for detection (Hansen et al., 2009; McKeefry, Murray, & Parry,

[1]Monash University, Melbourne, Victoria, Australia, [2]National Institutes for Quantum and Radiological Science and Technology, Chiba, Japan, [3]Tohoku University, Sendai, Japan, [4]National Institute of Information and Communications Technology (NICT), Suita, Japan, [5]Advanced Telecommunications Research Institute International, Kyoto, Japan

2007; Ayama, Sakurai, Carlander, Derefeldt, & Eriksson, 2004; Sakurai, Ayama, & Kumagai, 2003), although peripherally presented blue stimuli can increase in perceived saturation (Vanston & Crognale, 2018). This behavioral work is supported by neurophysiological studies showing that increasing retinal eccentricity is associated with decreasing density of cones (Curcio et al., 1990) and retinal ganglion cell density (Curcio & Allen, 1990) as well as a decline in cortical resource allocation (Daniel & Whitteridge, 1961). The sceptic's case is clear: The evidence indicates that peripheral color experiences are degraded or absent, and the naive introspection is unreliable.

Yet, support for the naive view of color experience across the visual field can be found in much of the same literature that sceptics cite. The same studies showing desaturation and hue-shifts with constant-size stimuli shown centrally and peripherally also show that if the stimuli are magnified in size as they increase in eccentricity, participants assign stimuli a consistent color (Hansen et al., 2009; Ayama et al., 2004; Sakurai et al., 2003; Abramov, Gordon, & Chan, 1991; Gordon & Abramov, 1977). There is also corresponding neurophysiological support for color experience across the visual field. Although cone densities decline dramatically with eccentricity, decreasing the resolution of incoming perceptual information, the retinal ganglion cells that pool photoreceptor inputs compensate for this with increasingly large receptive fields, ensuring color information is still available for downstream processing (Curcio & Allen, 1990). Similarly, the axonal projection zones of these retinal ganglion cells into the lateral geniculate nucleus and then to the primary visual cortex become larger with increasing eccentricity (Lennie, 1998). These neurophysiological findings do suggest that peripheral color perception is different from central perception in terms of acuity and contrast sensitivity, yet while also making plausible the naive case for color experiences across the visual field (Haun, 2021).

Resolving the conflict between these two positions requires a definition of what it would mean for color experiences to "feel the same" across the visual field. The majority of traditional color psychophysics research does not address this, instead focusing on related questions. One such example is "*under what conditions can the same physical stimulus induce different color experiences at different locations?*" (Figure 1A). Although important, it provides data only on the stimulus–experience relationship, rather than the experience–experience relationship, which we are after here. In addition to the material described above, there are data to suggest that the same stimulus can elicit a different stimulus–experience relationship even at different retinal locations of the same eccentricity (Afraz, Pashkam, & Cavanagh, 2010).

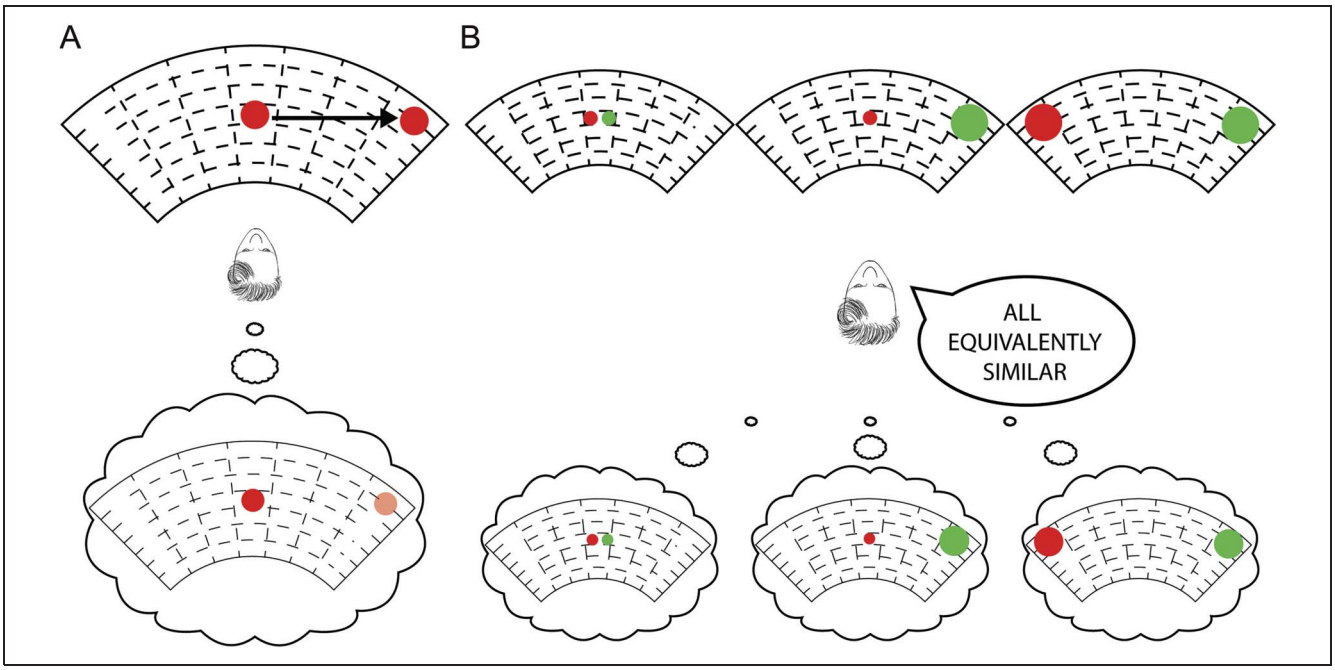

Figure 1. Traditional psychophysics and a novel equivalence test of color experience equivalence. There has long been controversy as to whether color experiences are "the same" across the visual field. The controversy originates from two issues: (1) how to measure color experience and (2) how to define "the same." In this article, we offer novel solutions to both of these issues. (A) A majority of traditional psychophysics has been concerned with whether "physically" identical stimuli generate the "same" color experience when presented at different retinal locations. These studies often used isolated instances of a single stimulus presented at a time, to which participants responded with "same/different responses." The results from this line of research tell us that small peripheral stimuli appear desaturated under certain conditions, demonstrating some "physically same" stimuli can be perceived as "different." (B) In contrast, our novel approach is concerned with whether "the structure of color experiences" differs at different visual field locations. To characterize the structure, we focus on the similarity relationships between a set of color experiences at different locations, with one possible pair out of the set shown here.

In contrast, we are seeking to answer "*whether color experience* ***relationships*** *are the same or not at different locations in the visual field*" (Figure 1B). We have previously suggested that characterization of an experience in relation to other possible experiences a participant could be having is sufficient to assess the notion of subjective equivalence that is meant by "the experiences felt the same" (Fink, Kob, & Lyre, 2021; Tsuchiya & Saigo, 2021). To operationalize the concept of "sameness" and characterize experience relationships, we will make use of "subjective equivalence tests." One possible equivalence test is the systematic pairwise comparison of the similarity between experiences at different visual locations. If comparing Experiences A or B to every other possible pairing of experiences results in sets of similarity judgments that are statistically indistinguishable, then A and B can be considered equivalent. This equivalence test will allow us to capture and compare aspects of experiences at different visual locations in a systematic manner. Conceptualizing experiences in this manner is becoming increasingly prominent in consciousness science (Fink et al., 2021). In addition, this relational approach is also related to representational similarity analysis of neural representations, a prominent method for linking brain activity to behavior (Kriegeskorte & Kievit, 2013). Here, we focus on the similarity relationships among color experiences in central and peripheral vision and test whether they are equivalent, so as to assess whether the experiences themselves differ across the visual field.

Partial inspiration for this approach comes from similarity experiments that have been used previously to build geometric models of color experience. The earliest publication was performed by Helm (1964), who asked participants to place colored chips at distances from other chips that were proportional to their subjective similarity. Unfortunately for the question at hand, no attempt was made to fix the retinal stimulus size or duration of presentation to preclude the possible eye movements. These considerations have also been neglected in more recent studies (Bonnardel et al., 2016; Burns & Shepp, 1988). Work using another paradigm to obtain subjective similarity between colors, asking participants to provide a numerical dissimilarity rating between two color patches, has also ignored the eccentricity of presentation. These studies have typically been performed with large stimuli (2–6 DVA), using both limited (e.g., 500 msec) and unlimited duration presentations (Bosten, Robinson, Jordan, & Mollon, 2005; Izmailov & Sokolov, 1991). Although this research provides support for a 3-D geometric model (e.g., HSV, CIELAB) of color experience in central vision, it cannot answer whether this model is valid for experiences across the visual field.

Furthermore, the previously described work examining color perception in the periphery (Ayama et al., 2004; Sakurai et al., 2003; Gordon & Abramov, 1977) cannot be used to assess whether the similarity relationships between color experiences changes across the visual field. This is because, in these studies, participants were forced to report their peripheral color experiences according to the 3-D model developed for relationships between central color experiences, without verifying that this model was appropriate for the periphery. There are two main shortcomings of this approach. The first is that participants were constrained to respond by clicking on points in a 3-D space. This may well have influenced participant responses, as the predefined space of possible responses already assumes the relationships between the colors. This constraint matters, as it has been shown that responses are different if participants are instructed to provide responses through the framework of color opponency compared with if they are uninstructed (Ennis & Zaidi, 2019). Second, it was not reported whether participants fully endorsed the responses that they provided or whether the responses provided merely had the lowest possible error given the reporting method's constraints. The relevance of this question is demonstrated by analogous experiments examining stimuli similarity under different illumination conditions. In these, participants are simultaneously able to report which achromatic color stimuli match "best" while denying that any possible stimuli adjustments can make the experiences equivalent when presented in different illumination contexts (Logvinenko & Maloney, 2006). This could not have been detected if the methods of the previous peripheral color experiments described above had been used. Addressing these issues requires comparing the similarity of color experiences at different eccentricities without prespecifying the color relationships.

We believe that a combination of three methods can be used to overcome the limitations of these prior studies. The first is to determine whether dissimilarity judgments for color experiences at different visual locations are correlated, which would indicate their structural equivalence. The second is to directly ask participants whether color experiences generated by identical stimuli presented at different locations are maximally similar (zero dissimilarity), which more directly establishes their equivalence. The third is to assess whether non-identical color pairs have the same similarity relationships at different locations, which establishes their experiential equivalence. If all of these are answered in the affirmative, it would establish the equivalence of the experiences at different locations.

This present study seeks to use these three methods to determine whether color experiences are equivalent across the visual field. We will do so by obtaining and analyzing the relationships between color experiences at varying eccentricities. Specifically, we will collect similarity ratings by pairwise comparison of experiences generated by stimuli presented either both in the central visual field (CC), one central and one peripheral (CP), or both peripheral (PP). To reduce the concern of possible eye movements as well as cognitive effects such as expectation, we will present two color patches for brief durations (250 msec) at unpredictable locations. Each participant

will rate the similarity between every combination of stimuli and eccentricity to allow for full determination of the color experience relationships with respect to the chosen set of stimuli. In addition, to relate our findings on color experience relationships to that of the previous work on perceptual changes with eccentricity, we will determine whether similarity judgments are affected by the size of the stimuli used to generate color experiences. In particular, we will examine whether keeping the size of the stimuli in the periphery results in these stimuli inducing altered color experience structures with respect to central vision. To buttress our novel methodology, we performed three control experiments described in Appendix D. They establish the sensitivity and reliability of our methods (e.g., robustness against display inhomogeneity). They also demonstrate that our methods can demonstrate the expected phenomenological consequences of a lack of S-cones at the central fovea (Magnussen, Spillmann, Stürzel, & Werner, 2004).

If the structure of dissimilarity relationships between color experiences has any degree of equivalence across the visual field, then participants should provide correlated dissimilarity judgments for sets of color stimuli shown at different eccentricities. For each participant, we will obtain the correlation between their judgments for each of the eccentricity conditions, that is, CC to CP, CC to PP, and CP to PP. We hypothesize that the mean across-participants correlation will be positive for each of these comparisons (H1A–C).

Correlation can establish one notion of structural equivalence, but given it is insensitive to uniform changes in judgments, it alone is insufficient to establish experiential equivalence across the visual field. Additional tests involve directly examining the magnitudes of the dissimilarity responses. For example, do identical color stimuli elicit completely similar experiences when presented at different visual field locations? If color experiences either do not exist, or are substantially degraded or altered in the periphery, then participants should provide the same distribution of responses for pairs of both physically identical and non-identical color pairs when one or both of the patches are presented at the periphery. Alternatively, if color experiences exist in the periphery and physically identical stimuli can elicit the same experience both centrally and peripherally, it should be possible to distinguish between the distribution of similarity responses given for identical versus non-identical stimuli pairs shown peripherally. For each participant, we will obtain the mean dissimilarity response for both physically identical and non-identical stimuli shown in each of the CC, CP, and PP conditions. The difference in the mean ratings for identical versus non-identical patches will be used to assess subjective equivalence as per the following hypotheses: Participants presented with peripherally magnified stimuli will provide a mean dissimilarity difference rating greater than zero (H2A–C) or alternatively close to zero, indicating no difference (H0).

Furthermore, we suspect that small stimuli shown in the periphery without peripheral magnification may not induce color experiences that are equivalent to when the same stimuli are presented centrally. To test this, we will use a 2×2 design altering stimulus size and peripheral magnification and observing the resulting mean ratings. We hypothesize that there will be an interaction effect on the normalized traces provided in the PP conditions (H2D).

Not only should identical stimuli elicit the same dissimilarity responses if color experience structures are equivalent across the visual field, but so too should any pair of color experiences, identical or non-identical, elicit the same dissimilarity response regardless of eccentricity. We will analyze whether this is indeed the case by examining whether participants make the same dissimilarity response for a color stimulus pair across the CC, CP, and PP trial conditions. We will quantify this by examining the between-conditions response variance for each color pair. If participants provide the same response for a given color pair under all three conditions, the variance in dissimilarity across those trials is zero. We will calculate variance in dissimilarity response between the three trial conditions across all color pairings for each participant and use the per-participant mean variance as the outcome measure. In contrast, if the similarity between color pairs is inconsistent between CC, CP, and PP trials, then the variance in dissimilarity across the three trials would be large. In this case, shuffling the correspondence between color-pair labels and a participant's similarity judgments for their CP and PP trials would have no effect on the mean variance, as their judgments would already be inconsistent across eccentricity conditions. We will examine if participants presented with peripherally magnified stimuli will provide consistent judgments across eccentricity conditions by assessing whether the difference in the mean variance of their judgments is greater than zero (H3), or whether they are inconsistent with no difference in mean variance between the original and scrambled judgments (H0).

All specific hypotheses are listed in the Hypothesis Summary Table (Table A1).

We have initially validated our methodology through online versions of both our preregistered experiment and replications of previously reported findings (Appendix D). After performing screen size and viewing distance calibrations, participants performed successive comparisons between color experiences generated by stimuli shown either centrally (1 DVA) or peripherally (10 DVA). Although the peripheral eccentricity was limited with respect to the maximal extent of human peripheral vision, this is still well outside the fovea, which comprises only a couple of degrees either side of the center of vision. For our preregistered experiments, we will expand on this pilot work by performing the same experiments but at a peripheral eccentricity of 60 DVA in a laboratory setting. All other aims and hypotheses remain the same.

METHODS

In this section, we provide the methods of the preregistered in-laboratory experiment. For our online pilot experiments, see Appendix D.

Ethics

All participants provided written informed consent before participating in the laboratory study, which has been approved by the Ethics Committee of the National Institute of Quantum Science and Technology. There are no conflicts of interest to declare. Participants were compensated for their time at a rate of ¥5000 JPY/day.

Design

Participants

Participants were obtained through responses to an e-mail list of registered participants held by the institute. Participants provided written consent before the commencement of the experiments.

Display Apparatus

Stimuli were presented on a Panoworks display (Orihalcon Technologies). Participants were seated 1359 mm away from the center of the screen, providing a screen size of 180 × 109.4 DVA (Figure A1). Stimuli were presented via Inquisit 6 (Millisecond) display software.

Stimuli

All stimuli were presented on a gray background corresponding to [0, 0, 0.5] in HSV color space. We used a stimulus set of nine fully saturated and luminant colors with maximal hue spacing (see Table B1 for full specification). All stimuli were presented as solid-colored circles that vary in position and size based on task conditions. All specifications of position are made with reference to the center of the screen (Figure 2A and B).

Centrally presented color stimuli were presented with their center at 1 DVA to avoid overlap. As the fovea extends to approximately 2 DVA and the macula to approximately 9 DVA, these stimuli are well within central vision. Peripheral stimuli were centered at 60 DVA, well outside of foveal vision (Figure 2B).

In experimental conditions with peripherally magnified stimuli, peripheral stimuli were scaled relative to central stimuli according to the average size of the V2 receptive field increase relative to central vision (Freeman & Simoncelli, 2011). For example, this accounts for a 4.5× increase in radius when presenting a stimulus at 10 DVA versus 1 DVA, or a 27× increase at 60 DVA. The diameter of the central stimuli was 0.1 DVA for the small stimulus and 2 DVA for the large stimulus. On any given trial, the stimuli are randomly selected to be centered anywhere

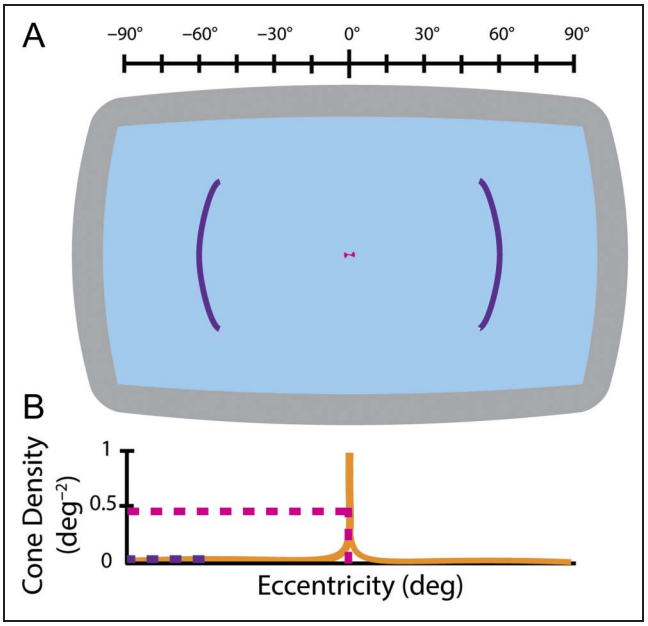

Figure 2. (A) Schematic of the achievable eccentricities of stimuli using our Panoworks display. Central and peripheral stimuli are presented at random locations centered on their appropriate eccentricity, which is shown schematically by the pink (1 DVA) and violet (60 DVA) lines, respectively. See Figure A1 for more details. (B) Cone photoreceptor densities at central and peripheral eccentricities relative to the retinal maximum for the range of the Panoworks display (Watson, 2014; Curcio et al., 1990). Eccentricities for (A) and (B) are shown by the scale above (A).

on the circle that corresponds to their appropriate eccentricity so long as the stimulus is not occluded by the edge of the screen (shown as colored areas in Figure 2A). Stimuli were always presented opposite each other (180° apart in polar angle).

Procedure

Instructions. After consent and screen size calibrations, participants were shown a sequence of task instructions (for the online pilot version, the demo is available here: https://mili2nd.co/76ub). Initially, they are presented with an animation of what the stimulus display will look like and instructed about what they will need to fixate on in the center of the screen. On the next page, they are informed that they will need to rate the similarity of the two stimuli presented on each trial and asked to ignore any differences in size and location between the stimuli. They were told they need to provide a judgment between 0 (most similar/least different) and 7 (least similar/most different) by clicking the appropriate integer, where the numbers are displayed in a circular fashion around the center of the screen after stimuli presentation (Figure 3). Participants were also encouraged to use the full range of possible values when providing responses. On this same screen, they are also presented with an animation of responses being provided on the response screen. After they have made their choice, they click on the center of

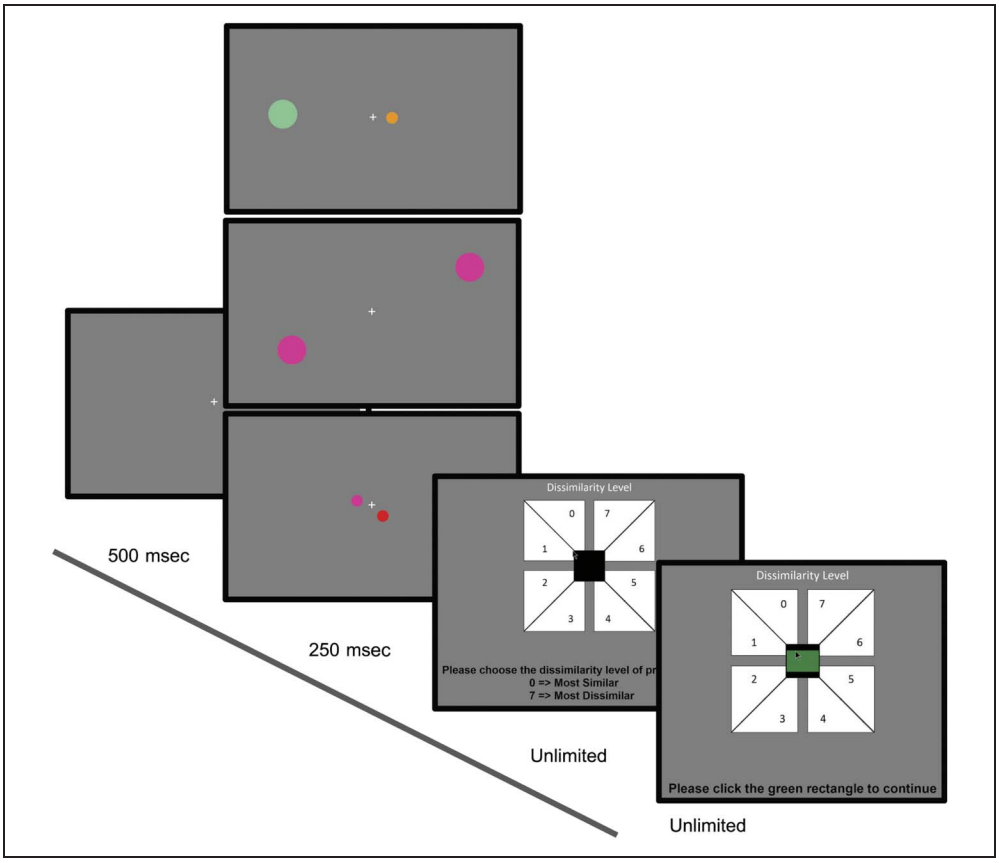

Figure 3. Schematic of the main trial procedure. Each trial begins with a 500-msec fixation cross. This is followed by stimuli presentation for 250 msec in either the CC (bottom), PP (middle), or CP (top) configuration. Participants are then asked to rate the dissimilarity of the two stimuli on the response screen by clicking on one of eight values. After responding, participants are asked to click on a green rectangle in the center of the screen to encourage fixation back to the center.

a rectangle in the center of the screen to direct their fixation and the mouse pointer back to the center for the next trial.

On the next screen, they are informed that some trials will be catch trials, where no stimuli will be displayed and, instead, participants are to select a specified value. Last, to provide context for similarity rating judgments, participants are also shown a 3 × 3 grid of colored circles corresponding to all the color stimuli they will see throughout the experiment (Figure B1). Subsequently, they perform nine practice trials to familiarize them selves with the experimental procedure. For these trials, they were provided with feedback on what selection they made, consisting of both the value they selected and the text "Very Similar," "Similar," "Different," or "Very Different" for selections of 0–1, 2–3, 4–5, 6–7, respectively. At the cessation of these practice trials, they were asked to press the SPACE button to proceed to the main trial set.

Main task. In each trial, participants first fixate on a central fixation cross for 500 msec. After this, there are three possible trial types. In main trials, two stimuli will be displayed, either in a center–center (CC), center–periphery (CP), or periphery–periphery (PP) configuration (Figure 3). The stimuli are presented for 250 msec, after which a response screen is displayed. By clicking the appropriate location, participants are asked to report the perceived similarity between the stimuli in terms of their color. In catch trials, where no color stimuli are displayed, participants are presented with a response screen and are directed to select a specified value (Figure B2). To initiate the next trial, they are directed to click a rectangle in the center of the screen.

Within each experimental condition, each participant completed 81 CC, 162 CP, and 81 PP main trials and 10 catch trials, equaling 334 trials in total per experimental condition. Trials are randomly ordered within sequence, and a different sequence will be presented to each participant. Each participant completed each of the four experimental conditions: large stimulus, peripheral magnification (MAG BIG); large stimulus, no peripheral magnification (FIX BIG); small stimulus, peripheral magnification (MAG SMALL); and small stimulus, no peripheral magnification (FIX SMALL). The sequence in which participants perform the experimental conditions was randomized across the different participants.

Sampling Plan

Participant Exclusion

Participant quality control occured both before and after the main experiment. Before the experiment, participants were excluded if they self-report visual acuity or color vision impairments, including participants who normally use glasses. After collecting the data and based on our online pilot experiments, participants who score < 77% on catch trials or who do not complete all main trials were

excluded (Figure C1; Appendix D). All other participants were included in the data analysis.

Bayes Factor Design Analysis

We use the Bayes Factor (BF) Design Analysis approach to determine the number of participants required to provide compelling evidence for or against our hypotheses (Table A1, Summary Table). We use a sequential design with maximum participant approach to recruit participants until either: the BF provides strong evidence for the null hypothesis ($BF < 0.1$) or the alternative hypothesis ($BF > 10$); or a total of 50 participants (without counting those that are excluded by the above criteria) has been reached (Schönbrodt & Wagenmakers, 2018). More specifically, the BF for each and all of Hypotheses 1A, 1B, 1C, 1D, and 2A must each exceed one of the thresholds for recruitment to be halted before reaching 50 participants. We also collect a minimum of five participants who pass the exclusion criteria. We compute BFs using the *bayestestR* package of the R Environment forStatistical Computing (Makowski, Ben-Shachar, & Lüdecke, 2019). BFs are calculated as appropriate for each hypothesis (see Analysis Plan section for further details).

Analysis Plan

Planned Analyses

Below, we use the results of the online pilot experiments to explain our analyses strategies and detailed hypotheses (Figures 4–7, Appendix D).

Descriptive statistics. For the registered in-laboratory experiment, we visually demonstrate the group average dissimilarity matrix as in Figure 4. We also perform metric multidimensional scaling (MDS) on the dissimilarity data to assess whether the data recapitulate the previously reported color hue rings. Last, for the planned experiments, we provide the within-subject response correlation across the first and second pass per condition. We did not collect this for the pilot study. In our pilot data, all groups except the participants shown small, unmagnified stimuli (Figure 4A and B) provide clearly structured responses across the full range of dissimilarities.

We have also collected $n = 1$ control participant data for one of the stimulus and magnification groups using the Panoworks display, confirming the feasibility of our study.

Are the overall structures of similarity reports equivalent for different eccentricities? (Hypothesis 1). If color experiences are equivalent across the visual field, then dissimilarity reports for color pairs experiences at one eccentricity should be correlated with those at another eccentricity (Figure 5A). For each participant, we obtain the correlation between their judgments for each of the eccentricity conditions, that is, CC to CP, CC to PP, and CP to PP (Figure 5B).

For the registered laboratory experiment, we hypothesize that **participants presented with peripherally magnified stimuli (at 60 DVA) will provide judgments for each eccentricity condition that are positively correlated with each other (H1A: CC to CP, H1B: CC to PP, H1C: CC to CP)**. We assess this by first Fisher-*Z* transforming the correlations and then performing a one-sample Bayesian *t* test with the default Jeffreys prior for the variance and Cauchy prior for the effect size for each hypothesis. We consider a $BF \geq 3$, indicating that the distribution of correlations is nonzero, to support our hypotheses. The online pilot data (periphery at 10 DVA) we have already collected support this hypothesis (Figure 5C, Table C1).

Are physically equivalent stimuli subjectively equivalent at different eccentricities? (Hypothesis 2). If color experiences are equivalent across the visual field, then an experience of a color should be reportable as identical to the same color presented elsewhere. Assuming this is the case, participants may report dissimilarity values for the comparison of physically identical stimuli that are close to zero and significantly smaller than for non-identical stimuli pairs. We assess whether dissimilarity values for identical pairs are significantly smaller than for non-identical pairs. For each participant, we obtain the mean dissimilarity response to physically identical stimuli, termed a "normalized trace," in each of the CC, CP, and PP conditions (Figure 6A and B), as well as their mean rating for non-identical stimuli (Figure 6C). The per-participant difference between their mean response to non-identical pairs and the normalized trace is used to assess subjective equivalence, termed a "normalized trace difference" (Figure 6D).

For the registered laboratory experiment, we hypothesize that **participants presented with peripherally magnified stimuli (at 60 DVA) will provide a normalized trace difference significantly greater than zero for each of the CC (H2A), CP (H2B), and PP (H2C) conditions**. We assess this by fitting the normalized trace difference of these participants using a Bayesian beta regression with no regressors. A beta regression is appropriate as the normalized traces are bounded by the minimum and maximum possible dissimilarity responses (Cribari-Neto & Zeileis, 2010). We perform a transformation to ensure boundary offset (Smithson & Verkuilen, 2006). The data are fitted using the R package Rstanarm (Goodrich, Gabry, Ali, & Brilleman, 2020). We use the default prior that the data mean (model intercept) will be centered around zero, which corresponds to a prior belief that the participants provide the same responses to non-identical and identical color pairs. We report the intercept estimate and its Bayesian 95% highest-posterior-density (HPD) interval, as well as the BF_{10} for the intercept

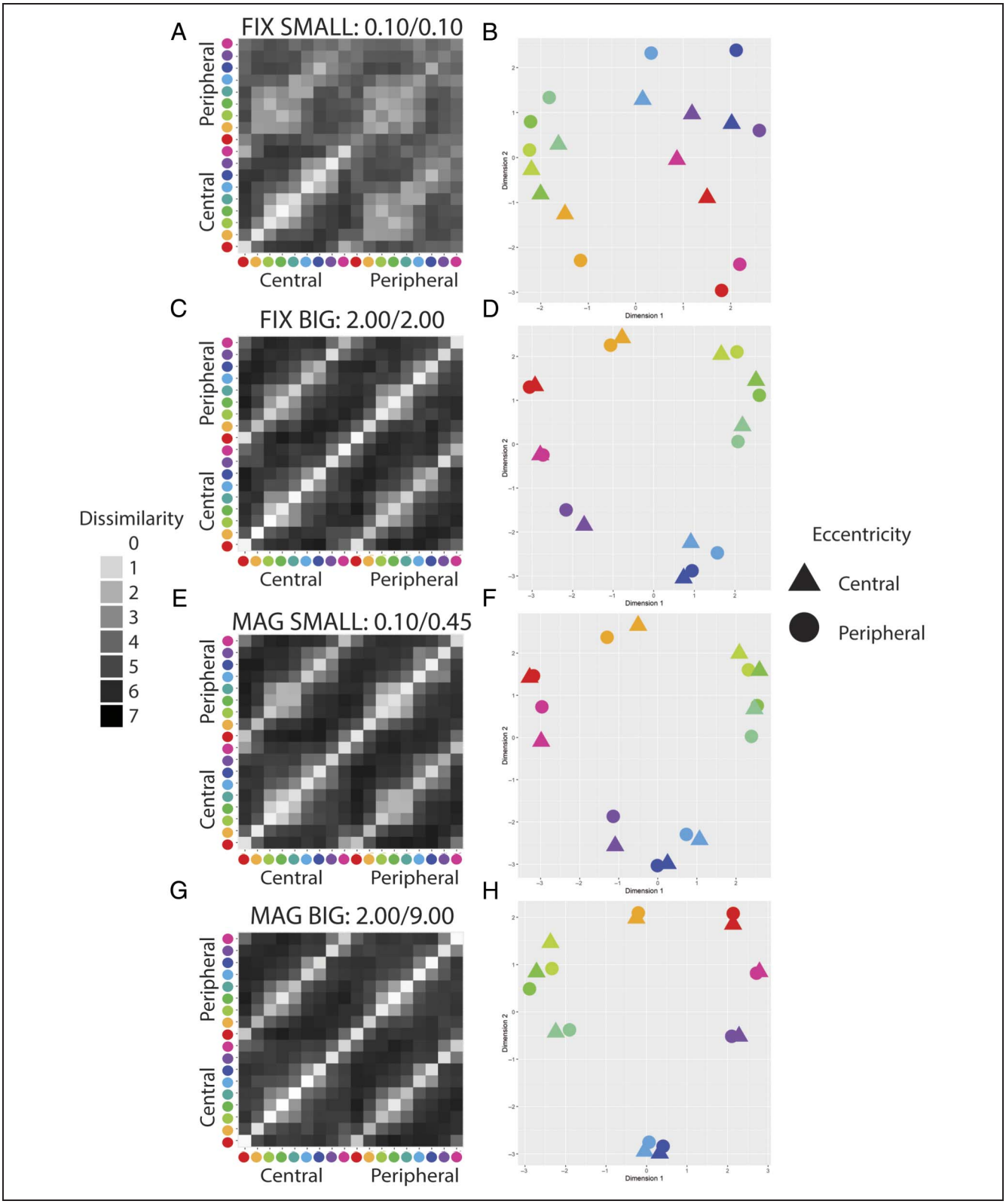

Figure 4. Results of the pilot online experiment (n = 13–17). (A, C, E, G) Group mean dissimilarity matrices for each experimental condition. The data have been symmetrized across the main diagonal. (B, D, F, H) The corresponding MDS plot for each condition. Circles and triangles indicate centrally and peripherally presented stimuli, respectively. Peripheral stimuli are centered at 10 DVA. FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×4.5) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

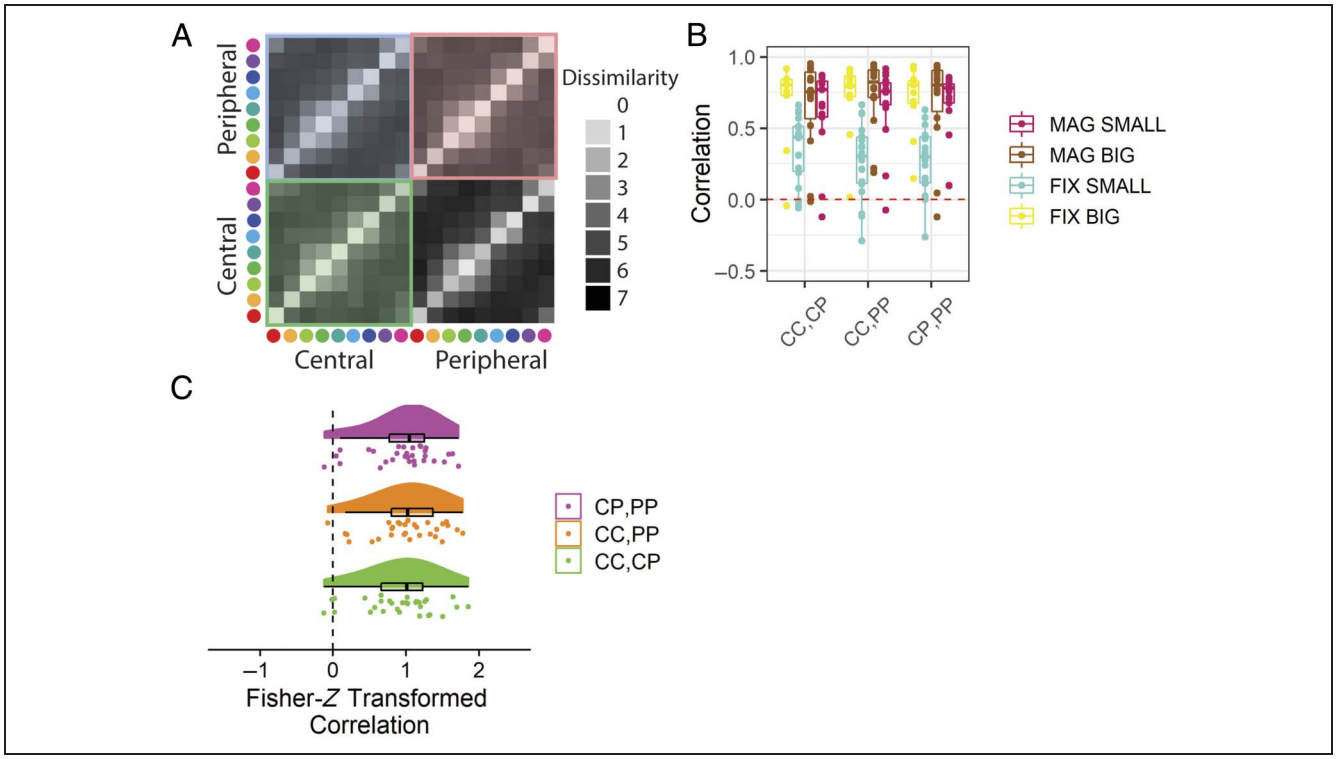

Figure 5. Results from the pilot online experiment that address whether similarity reports for the same stimuli shown at different eccentricities are correlated. (A) A schematic illustration with the set of dissimilarity values reported with the values being compared marked with colored shading. (B) The within-subject correlations between judgments given in different eccentricities. The red dashed line indicates zero correlation. (C) The Fisher-*Z* transformed within-subject correlations between judgments across eccentricity for participants shown peripherally magnified stimuli. Peripheral stimuli are centered at 10 DVA. Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. CC = central–central; CP = central–peripheral; PP = peripheral–peripheral; FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×4.5) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

using the Savage-Dickey density ratio (Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010). We consider a BF ≥ 3, indicating that the model intercept is less than zero, to support our hypothesis. The online pilot data (periphery at 10 DVA) we have already collected support this hypothesis (Figure 6E, Table C2).

We suspect that **small stimuli shown in the periphery may not induce color experiences that are equivalent to when the same stimuli are presented centrally. Furthermore, we expect that this can be corrected by magnifying the size of stimuli in the periphery (H2D)**. We assess this by fitting all the participant normalized trace data with a Bayesian beta regression with stimulus size, peripheral magnification, and Size × Magnification interaction as regressors. We use the default prior, which corresponds to a prior belief that the coefficient for all of the regressors is zero. We consider a BF ≥ 3 for the interaction term to support our hypothesis. The online data we have already collected are inconclusive for this hypothesis (Figure 6F; Table C3).

Are color-pair comparisons equivalent across the visual field? (Hypothesis 3). If color experience structures are equivalent across the visual field, any pair of color experiences should elicit the same dissimilarity response regardless of eccentricity. We analyze whether this is indeed the case by examining whether participants make the same dissimilarity response for a color stimulus pair in the CC, CP, and PP trial conditions. If participants provide the same responses for all three conditions, the variance in dissimilarity across those trials is zero (Figure 7A and B). In contrast, if the response to a color pair in a particular trial type is uninformative with respect to the other trial types (i.e., the correspondence in dissimilarity values for a particular color pair under CC, CP, and PP conditions is no greater than between color pair dissimilarity values randomly sampled from a participant's judgments), then the variance in dissimilarity across trials in the three conditions would be substantial (Figure 7B). We assess whether participants provide consistent responses across eccentricity conditions by comparing the mean variance of their actual color pair judgments across CC, CP, and PP conditions to the case where the corresponding CP and PP judgments are randomly scrambled (Figure 7B; Figure C3). If the difference in mean variance between participants' original and scrambled data is greater than zero, then participants are providing consistent responses across eccentricity conditions. Thus, we will obtain the difference in mean variance for original and CP/PP-scrambled dissimilarity responses across the three

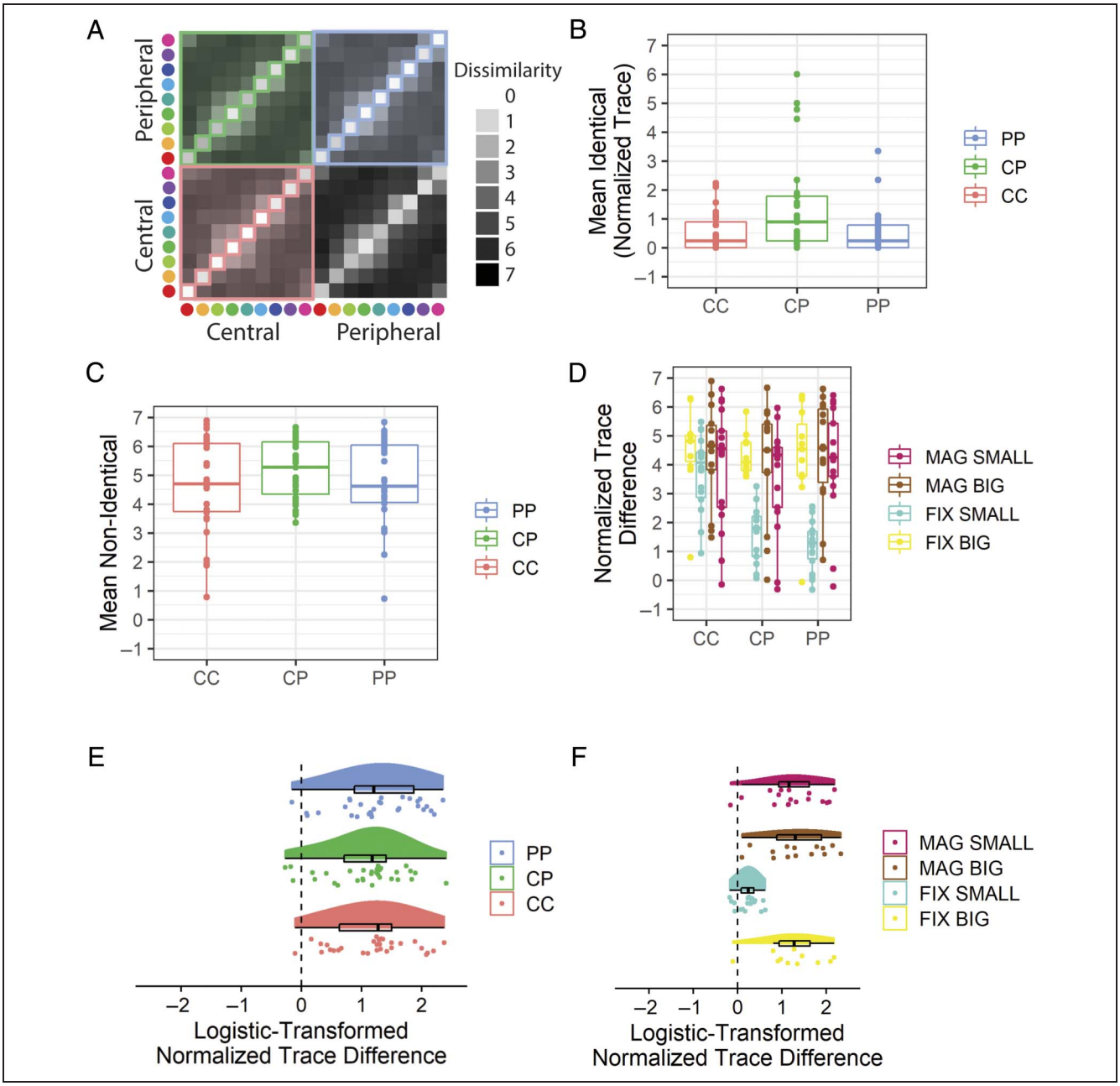

Figure 6. Results from the pilot online experiment that address whether identical color stimuli are reported as the same at different viewing locations. (A) A schematic illustration with the set of dissimilarity values reported for identical color stimuli in each condition marked by the diagonal squares, whereas non-identical color stimuli are marked by colored shading. (B) The mean rating given to identical color pair stimuli in each eccentricity condition, for participants shown peripherally magnified stimuli. (C) As per (B), but for non-identical color pairs. (D) The difference between the mean dissimilarity value for non-identical and identical color patch pairs, called a "normalized trace difference." The normalized trace difference is plotted for each condition for each participant. (E) The normalized trace difference values for each condition from participants shown peripherally magnified stimuli. The values are scaled from 0 to 1 based on the maximum and minimum normalized trace differences possible and then logistically transformed. (F) The transformed normalized trace differences for the PP condition for all participants. Peripheral stimuli are centered at 10 DVA. Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. CC = central–central; CP = central–peripheral; PP = peripheral–peripheral; FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×4.5) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

trial conditions and all color pairings for each participant (Figure 7C and D).

We hypothesize that participants presented with peripherally magnified stimuli will provide a difference in mean variance for color pairs shown in different eccentricities significantly greater than that expected from inconsistent color pair judgment responses (H3). We assess this by fitting the

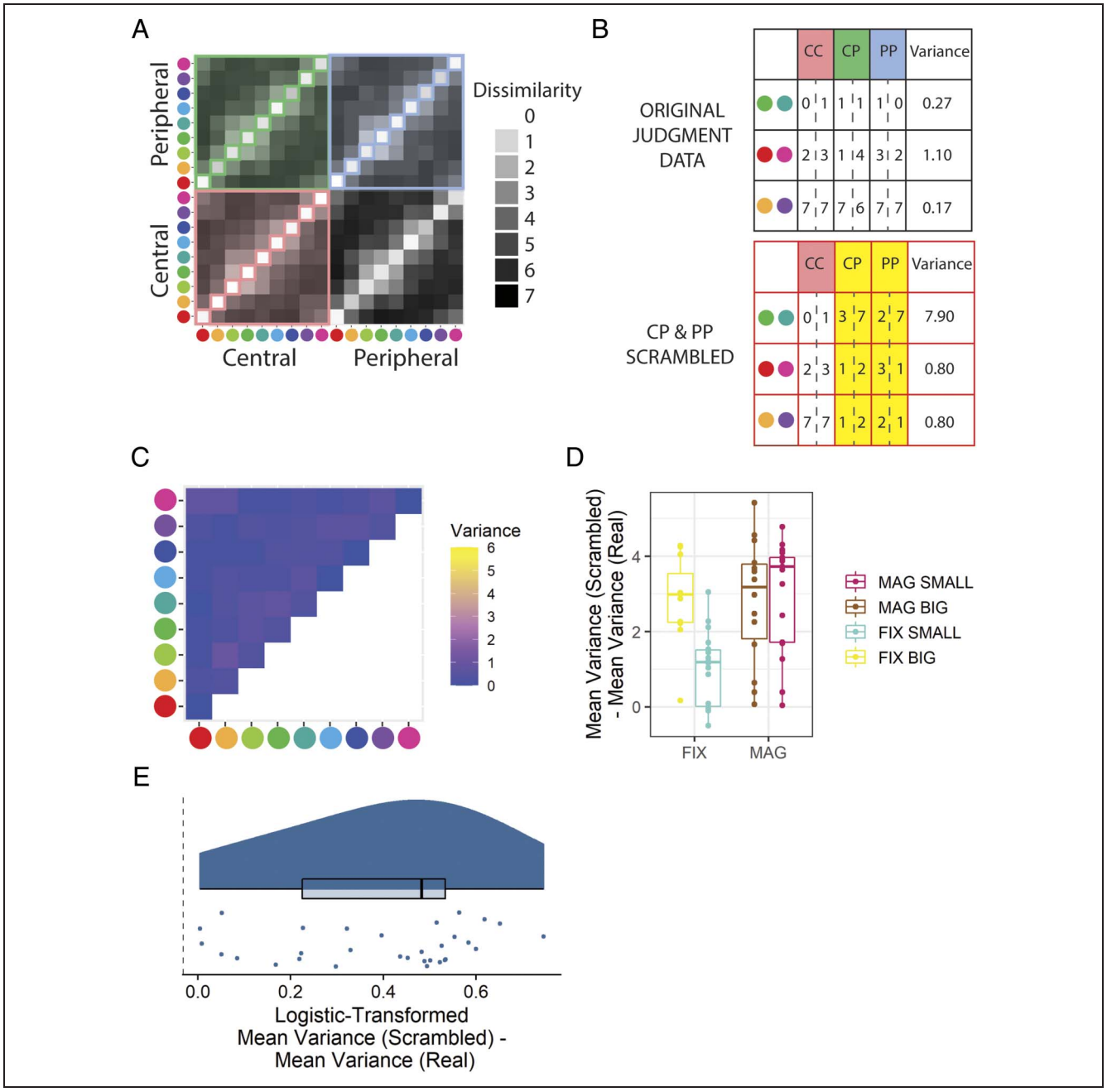

Figure 7. Results from the online experiment that address whether arbitrary pairs of color stimuli are reported to have the same dissimilarity at different viewing locations (n = 13–17). (A) The set of dissimilarity values collected from participants for each possible color pair in each condition is marked by the colored shading. Participants provided two judgments for each non-identical color pair in each condition. (B) A demonstration of how variance in dissimilarity values is calculated for a given color pair shown in different conditions. Two judgments are made for each color pair in each condition, meaning the variance is calculated across six values. In the analysis, this variance is compared with the mean variance that would be seen if random color pairs were sampled as opposed to the same color pair under different eccentricity conditions. See also Figure C3. (C) The mean variance across participants shown peripherally magnified stimuli for each color pair. (D) The difference in mean color-pair variance for scrambled and real ratings provided by each participant. (E) The difference in mean color-pair variance between scrambled and real data for each participant shown peripherally magnified stimuli. The values are scaled from 0 to 1 based on the maximum and minimum mean variance difference possible and then logistically transformed. Peripheral stimuli are centered at 10 DVA. Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. CC = central–central; CP = central–peripheral; PP = peripheral–peripheral; FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×4.5) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

mean variance difference from these participants using a Bayesian beta regression with no regressors and the default prior that the data mean (model intercept) will correspond to 0. We consider a BF ≥ 3, indicating that the mean (model intercept) is greater than that expected from inconsistent responding, to support our hypothesis. The online data we have already collected support this hypothesis (Figure 7E; Table C4).

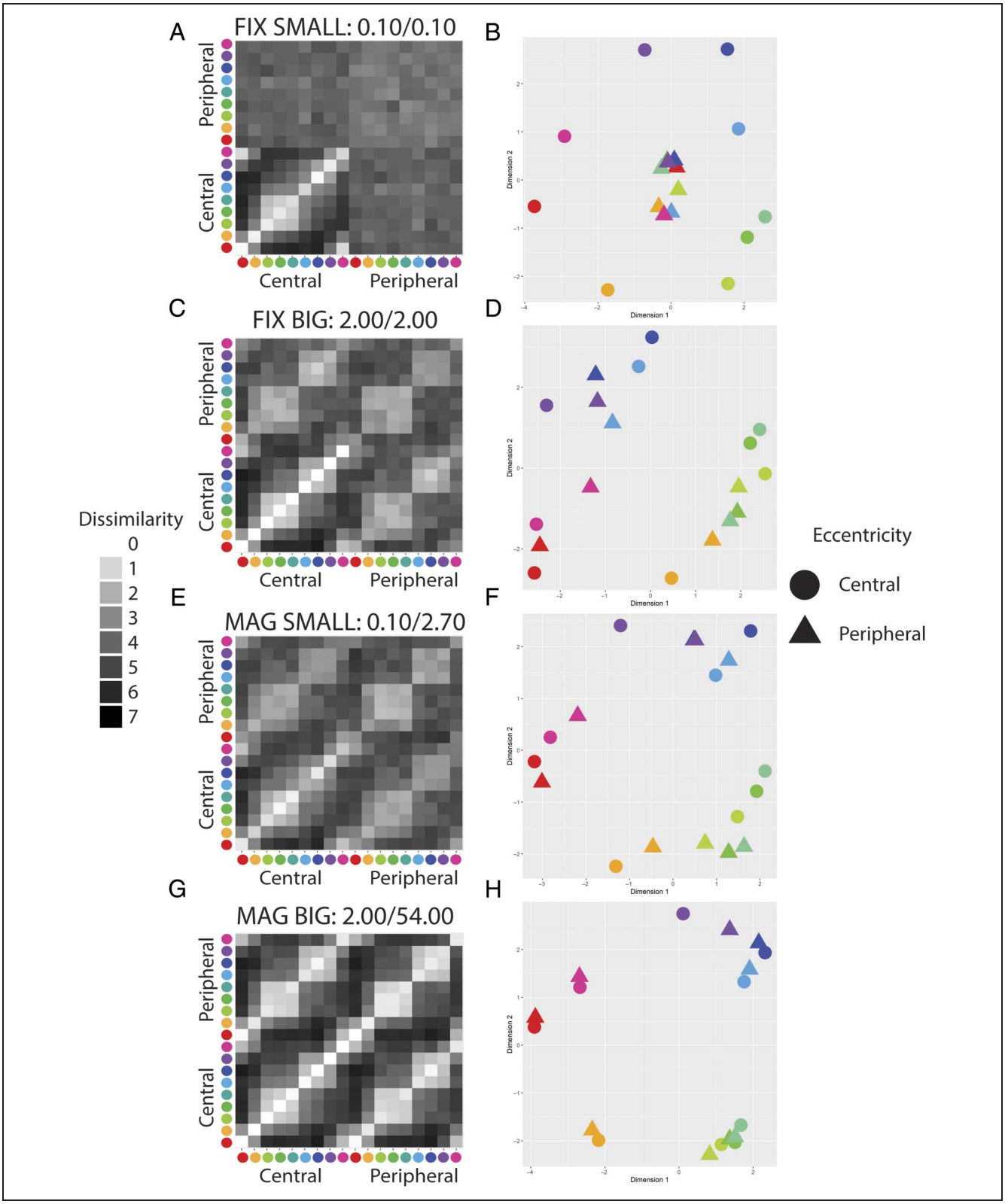

Figure 8. Visualization of the participant dissimilarity judgments for the in-laboratory experiment (n = 14). (A, C, E, G) Group mean dissimilarity matrices for each experimental condition. The data have been symmetrized across the main diagonal. (B, D, F, H) The corresponding MDS plot for each condition. Circles and triangles indicate centrally and peripherally presented stimuli, respectively. Peripheral stimuli are centered at 60 DVA. FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×27) stimuli;. SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

Table 1. A Summary of the Stimulus Eccentricity and Size Parameters across the Online and In-laboratory Experiments

Name	*Central Stimulus Eccentricity (DVA)*	*Central Stimulus Diameter (DVA)*	*Peripheral Stimulus Eccentricity (DVA)*	*Peripheral Stimulus Diameter (DVA)*
Online				
FIX SMALL	1	0.10	10	0.10
FIX BIG	1	2.00	10	2.00
MAG SMALL	1	0.10	10	0.45
MAG BIG	1	2.00	10	9.00
In-laboratory				
FIX SMALL	1	0.10	60	0.10
FIX BIG	1	2.00	60	2.00
MAG SMALL	1	0.10	60	2.70
MAG BIG	1	2.00	60	54.00

RESULTS

Participants

For the in-laboratory experiment, we recruited 14 healthy volunteers (8 women, mean age = 36.4, *SD* = 13.8) with normal or corrected-to-normal vision (color vision not formally assessed) before reaching our stopping criteria. No participants were rejected because of meeting exclusion criteria.

Descriptive Statistics

Informal visual inspection of both the group mean participant dissimilarity matrix and a 2-D metric MDS representation of that data for each experimental condition suggests that whether color experiences are equivalent across the visual field depends on the size of the stimuli that elicit them (Figure 8, Table 1). Photographs of the monitor display for these conditions can be seen in

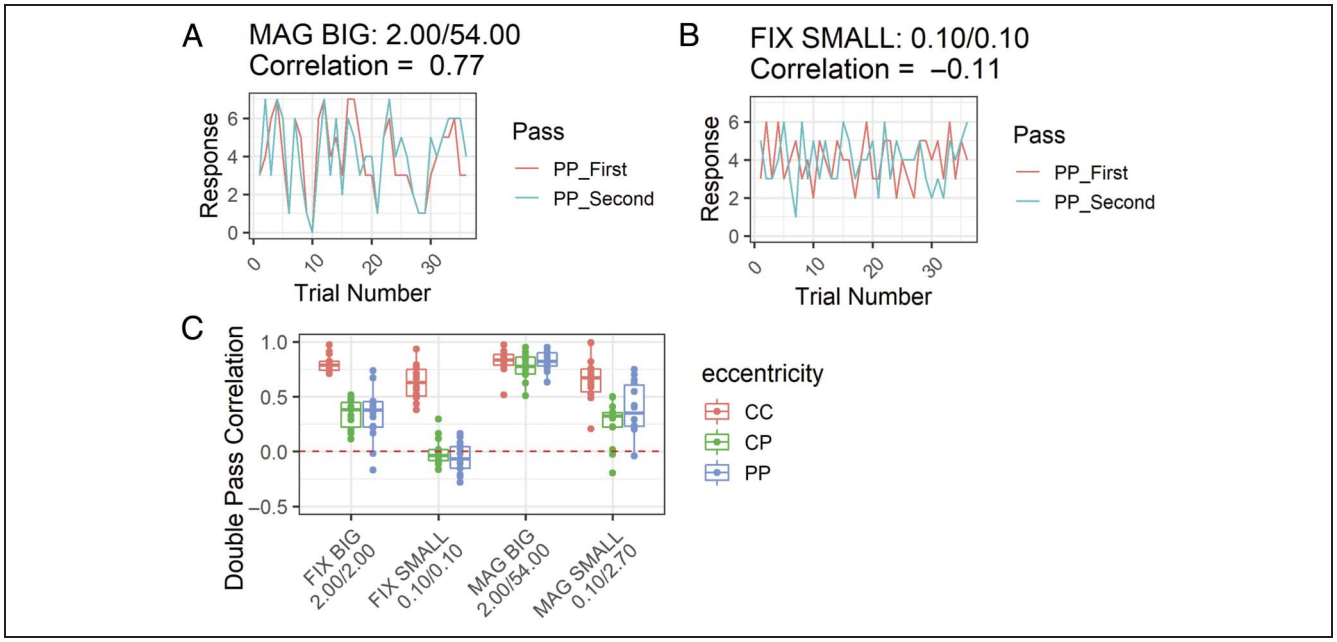

Figure 9. Results from the double-pass analysis for judgment consistency. (A) The first and second responses for each color pair in the PP presentations for Participant 1 in the MAG BIG condition. The high correlation between the two passess indicates consistent responding. (B) As per (A) but for Participant 1 during the FIX SMALL condition. Low correlation indicates inconsistent responding, perhaps because of failure to perceive the stimuli. (C) The double pass correlation across all participants for each condition and presentation type. FIX = peripheral stimuli of fixed size with respect to central stimuli. Peripheral stimuli are centered at 60 DVA. MAG = peripherally magnified (×27) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA; CC = central–central; CP = central–peripheral; PP = peripheral–peripheral comparison.

Figure A2. Raw responses for the individual participants can be seen in Figure A3.

Each participant performed each color pair judgment twice for each eccentricity condition within each experimental condition, with the "first pass" and "second pass" together comprising a "double pass paradigm." Determining the correlation between a participant's responses for their first and second pass allows us to judge how consistent they are in their responses given identical stimulus conditions. When participants can see the stimuli in a consistent manner, the correlation should be high (Figure 9A). Note that high correlations are possible even with differences in the quality of color experiences at the fovea and periphery (e.g., if peripheral stimuli look desaturated in a consistent manner, it should still result in high correlations; Decock & Douven, 2013). When participants cannot see the stimuli per se, or at least not their colored aspect, their similarity responses may become more random, resulting in lower correlations across passes (Figure 9B). Resultantly, examination of the double pass correlations across all experimental conditions for all participants gives some suggestion of their ability to perceive the stimuli (Figure 9C). In the MAG BIG condition, participants give highly correlated judgments regardless of eccentricity, suggesting consistent perception of the stimuli. In contrast, participant judgments are uncorrelated for the CP and PP trials during the FIX SMALL condition, indicating a potential failure to perceive these small, unmagnified stimuli (either entirely or at least their color aspect). This matches the informal reports some participants spontaneously provided to the experimenter during the FIX SMALL condition claiming they could not see the color or location of the stimulus. Whereas CC judgments remain highly correlated in both the FIX BIG and MAG SMALL conditions, participant responses for CP and PP trials have a relatively reduced but still positive correlation.

Similarity Structure Correlations (Hypothesis 1)

We hypothesized that participants presented with peripherally magnified stimuli would provide dissimilarity judgments at each eccentricity condition that are positively correlated with each other (H1A: CC to CP, H1B: CC to PP, H1C: CC to CP). All of these hypotheses were supported (Figure 10C, Table 2).

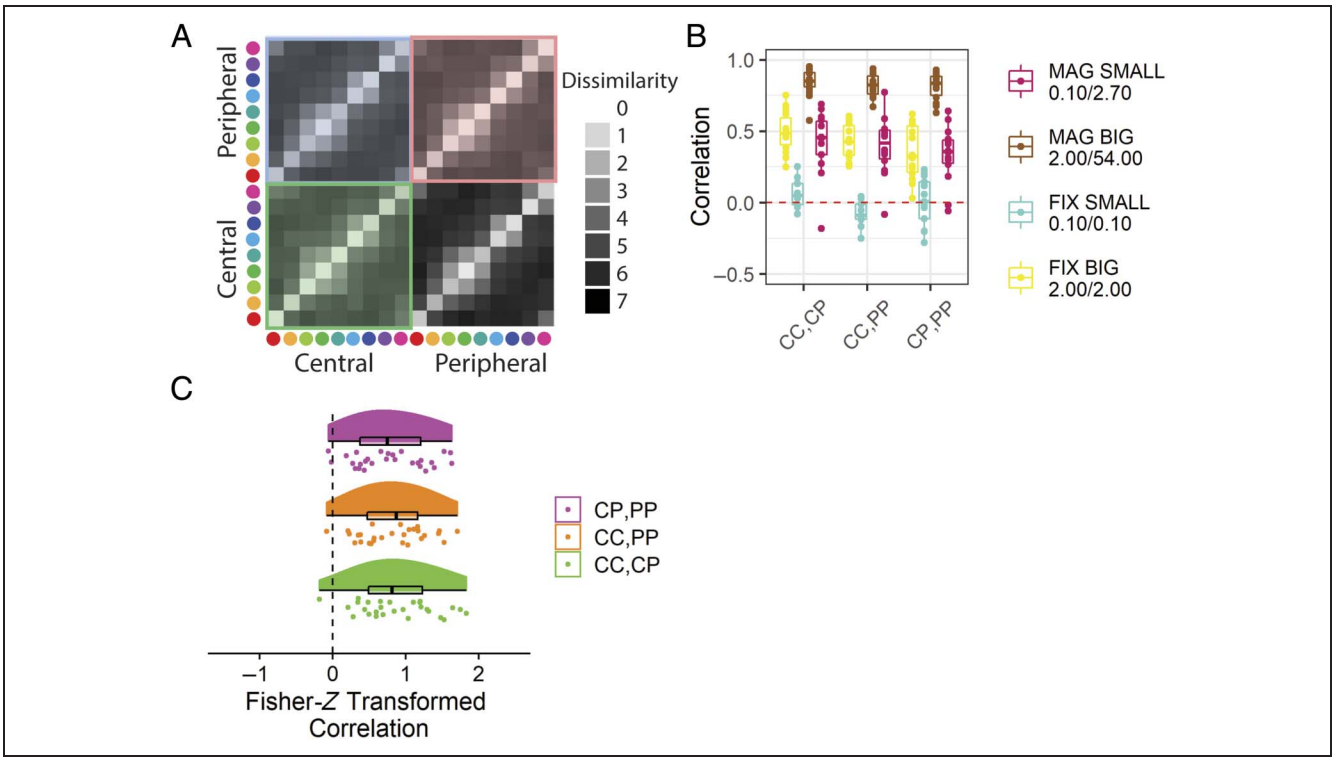

Figure 10. Results from the in-laboratory experiment that address whether similarity reports for the same stimuli shown at different eccentricities are correlated. (A) A schematic illustration with the set of dissimilarity values reported with the values being compared marked with colored shading. (B) The within-subject correlations between judgments given in different eccentricities. The red dashed line indicates zero correlation. (C) The Fisher-*Z* transformed within-subject correlations between judgments across eccentricity for participants shown peripherally magnified stimuli. Peripheral stimuli are centered at 60 DVA. Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. CC = central–central; CP = central–peripheral; PP = peripheral–peripheral; FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×27) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

Table 2. A Summary of the Experimental Results

	Similarity Structure Correlations		
Hypothesis	*Comparison*	*Mean Correlation*	BF_{10}
H1A	CC, CP	0.71	$> 10^9$
H1B	CC, PP	0.68	$> 10^7$
H1C	CP, PP	0.65	$> 10^7$

	Experiential Equivalence of Identical Stimuli			
Hypothesis	*Condition*	*Mean*	*95% HPD*	BF_{10}
H2A	CC	3.19	2.48 to 3.80	$> 10^5$
H2B	CP	2.21	1.45 to 2.91	$> 10^2$
H2C	PP	2.49	1.76 to 3.19	$> 10^4$

	Parameter	*Estimate*	*95% HPD*	BF_{10}
	Intercept	−0.10	−0.36 to 0.13	
	Size	0.5	0.12 to 0.83	
	Magnification	0.0	0.01 to 0.04	
H2D	Size × Magnification	0.0	−0.02 to 0.02	0.05

	Overall Judgment Consistency		
Hypothesis	*Mean*	*95% HPD*	BF_{10}
H3	2.13	0.97 to 3.23	11.17

The corresponding hypotheses are listed in the text and summarized in Table A1.

Experiential Equivalence of Identical Stimuli (Hypothesis 2)

We hypothesized that participants presented with peripherally magnified stimuli would provide a normalized trace difference significantly greater than zero for each of the CC (H2A), CP (H2B), and PP (H2C) conditions. All of these hypotheses were supported (Figure 11E, Table 2).

In addition, we hypothesized that there would be an interaction between peripheral magnification and stimulus size on participant dissimilarity judgments. This hypothesis was not supported (Figure 11F; Table 2).

Overall Judgment Consistency (Hypothesis 3)

We hypothesized that participants presented with peripherally magnified stimuli would provide a difference in mean variance for color pairs shown in different eccentricities significantly greater than that expected from inconsistent color pair judgment responses (H3). This hypothesis was supported (Figure 12E; Table 1).

DISCUSSION

We set out to examine whether color experience relationships are the same across the visual field, in contrast to previous psychophysics studies examining whether the same stimulus always evokes the same experience irrespective of retinal location. Our first hypothesis was that color judgments for color-pair experiences at different locations would be correlated, indicating structural equivalence. We found that this was indeed the case (H1A–C). Our second hypothesis was that (peripherally magnified) identical stimuli would evoke maximally similar (zero dissimilarity) experiences, supporting structural equivalence that goes beyond mere correlation. This too was supported by the data (H2A–C). Our third hypothesis was that non-identical color stimuli pairs would elicit the same similarity responses irrespective of presentation location, establishing their experiential equivalence at different locations. This was also affirmed (H3). Moreover, participants could provide these responses in a consistent manner, as demonstrated by the double-pass analysis. Together, the collected data unambiguously demonstrate the structural

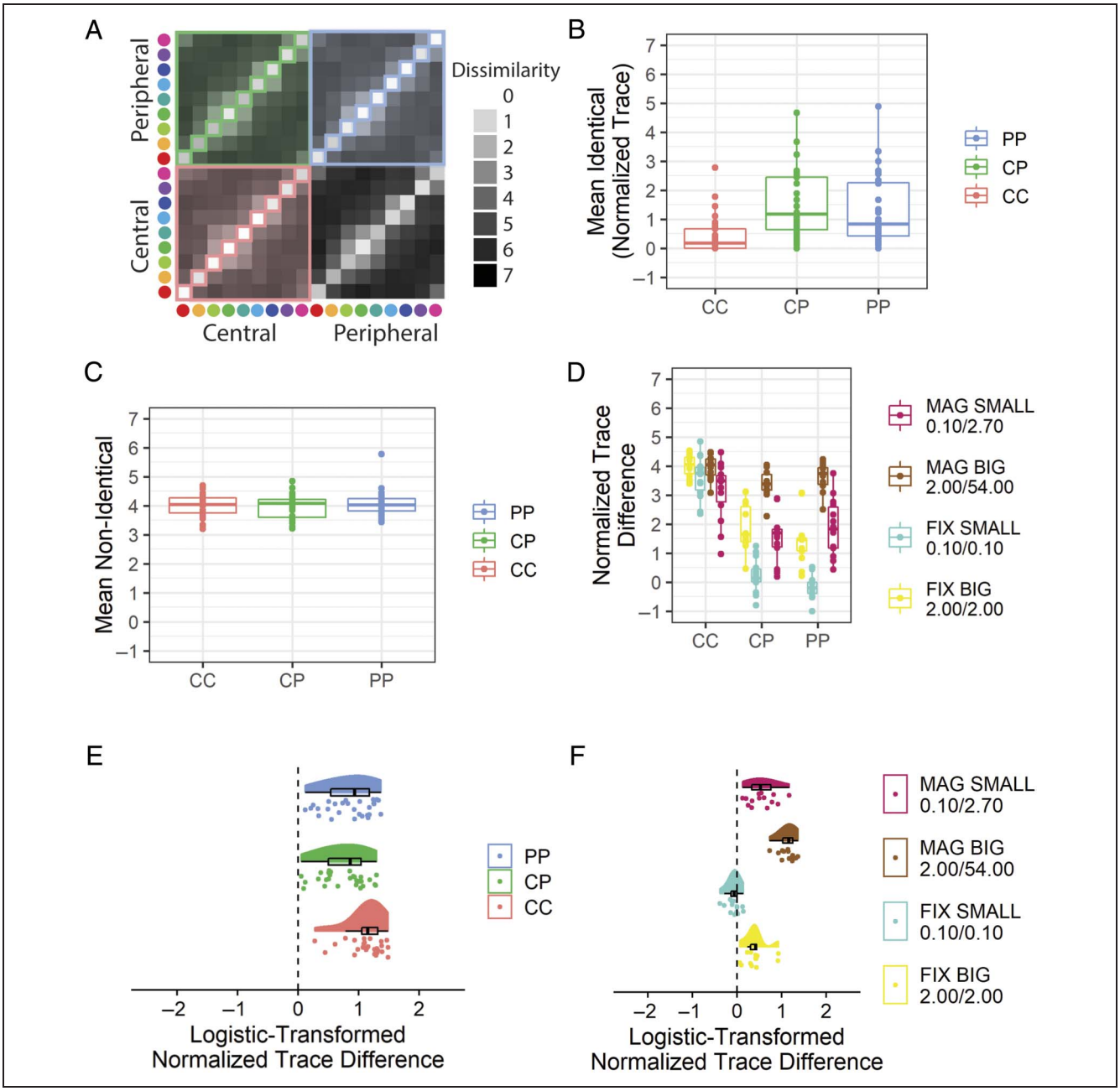

Figure 11. Results from the in-laboratory experiment that address whether identical color stimuli are reported as the same at different viewing locations. (A) A schematic illustration with the set of dissimilarity values reported for identical color stimuli in each condition marked by the diagonal squares, whereas non-identical color stimuli are marked by colored shading. (B) The mean rating given to identical color pair stimuli in each eccentricity condition, for participants shown peripherally magnified stimuli. (C) As per (B), but for non-identical color pairs. (D) The difference between the mean dissimilarity value for non-identical and identical color patch pairs, called a "normalized trace difference." The normalized trace difference is plotted for each condition for each participant. (E) The normalized trace difference values for each condition from participants shown peripherally magnified stimuli. The values are scaled from 0 to 1 based on the maximum and minimum normalized trace differences possible and then logistically transformed. (F) The transformed normalized trace differences for the PP condition for all participants. Peripheral stimuli are centered at 60 DVA. Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. CC = central–central; CP = central–peripheral; PP = peripheral–peripheral; FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×27) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

equivalence of color experiences across the visual field when elicited by peripherally magnified stimuli.

Our data-supported hypotheses in turn support the naive view that neurotypical humans perceive and experience color across their entire visual field. Peripheral color perception, that is, the ability for large peripheral stimuli to be consistently detected, appraised and compared, is demonstrated explicitly by the formal hypotheses

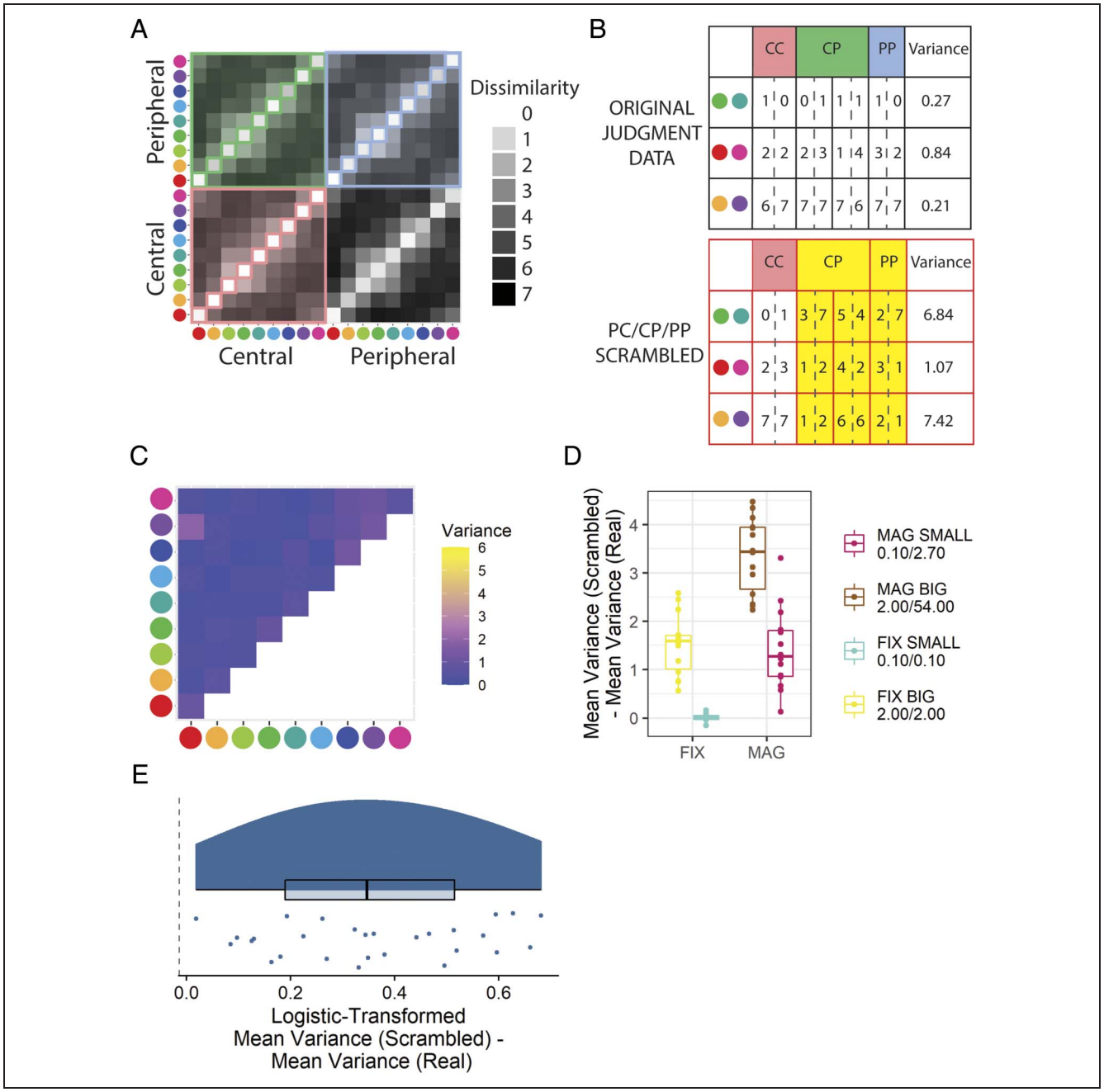

Figure 12. Results from the in-laboratory experiment that address whether arbitrary pairs of color stimuli are reported to have the same dissimilarity at different viewing locations. (A) The set of dissimilarity values collected from participants for each possible color pair in each condition is marked by the colored shading. Participants provided two judgments for each non-identical color pair in each condition. (B) A demonstration of how variance in dissimilarity values is calculated for a given color pair shown in different conditions. Two judgments are made for each color pair in the CC and PP conditions and four for each color pair in the CP condition, meaning the variance is calculated across eight values. In the analysis, this variance is compared with the mean variance that would be seen if random color pairs were sampled as opposed to the same color pair under different eccentricity conditions. See also Figure C3. (C) The mean variance across participants shown peripherally magnified stimuli for each color pair. (D) The difference in mean color-pair variance for scrambled and real ratings provided by each participant. (E) The difference in mean color-pair variance between scrambled and real data for each participant shown peripherally magnified stimuli. The values are scaled from 0 to 1 based on the maximum and minimum mean variance difference possible and then logistically transformed. Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. Peripheral stimuli are centered at 60 DVA. CC = central–central; CP = central–peripheral; PP = peripheral–peripheral; FIX = peripheral stimuli of fixed-size with respect to central stimuli; MAG = peripherally magnified (×27) stimuli; SMALL = central stimulus diameter of 0.1 DVA; BIG = central stimulus diameter of 2.0 DVA.

along with the double-pass correlations. In addition, peripheral color experiences, that is, "what-it-feels-like to experience colors in the periphery," were demonstrated to be structurally equivalent to those in central vision. Generalizing these findings outside the laboratory would suggest, as naively expected, that we really can perceive and experience a canopy of leaves as green, the clouds as gray, and the sky as blue in our peripheral vision.

However, we did not hypothesize, and our data do not suggest, that we perceive small stimuli equivalently when presented at different visual field locations. Our sole invalidated hypothesis was that there would be an interaction effect between stimulus size and peripheral magnification on dissimilarity responses for identical versus non-identical stimuli shown in the periphery. We formulated this hypothesis as we believed that only small, unmagnified, peripherally presented stimuli would fail to produce differing judgments for identical versus non-identical stimuli. We expected that either increasing stimulus size (from 0.10- to 2.00-DVA diameter) or peripherally magnifying the stimulus (from 0.10- to 2.70-DVA diameter) would fully attenuate this effect and produce equivalent judgments to when the stimuli were presented centrally. Instead of an interaction though, these judgments instead appeared to be affected additively by stimulus size and peripheral magnification. This suggests that for small stimuli to elicit fully equivalent color experiences at 60 DVA to those seen in the fovea, they need to be peripherally magnified by a greater factor than the V2 receptive field scaling that we chose. V4 scaling may potentially be a more appropriate choice (Freeman & Simoncelli, 2011).

There are at least two distinct phenomenological situations that may be occurring in the small, unmagnified stimuli paradigm. One possibility is that these stimuli were perceived and elicited color experiences, but with a genuine change in the experienced color and their corresponding similarities. For example, previous studies have suggested that peripheral stimuli appear less saturated (Sakurai et al., 2003; Gordon & Abramov, 1977), although see Rajananda, Peters, Lau, and Odegaard (2017). However, our double-pass analysis (Figure 9) revealed that repeated presentation of small 0.10-DVA stimuli pairs at 60 DVA produced entirely uncorrelated responses, whereas ~2.00-DVA stimuli responses were positively correlated yet reduced compared with larger stimuli. Furthermore, participants' responses to small stimuli pairs were inconsistent compared with larger stimuli, rather than of equal variance but different value. This pattern of results seems hard to reconcile with an explanation proposing participants perceived peripheral stimuli as uniformly and reliably less saturated in this experiment. The more likely scenario is that these small stimuli were simply harder to consciously detect than their larger counterparts, resulting in a failure to elicit reliable similarity responses.

Given our current findings of color perception in peripheral vision and related earlier work in the literature, why then do some claim that "the feeling that our entire 'subjective visual world' is richly colored…must be an illusion" (Chater, 2018)? We believe it stems from a particular interpretation of the known findings on how perceptual abilities differ across the visual field. Acuity for detecting stimuli of a fixed size drops with eccentricity (Strasburger et al., 2011), and this effect is more pronounced for chromatic than achromatic stimuli (Anderson, Mullen, & Hess, 1991; Mullen, 1991). Contrast-sensitivity for high spatial frequencies declines exponentially with eccentricity (Yang, Qi, & Makous, 1995), whereas in general, increased contrast is required for chromatic stimulus detection (Hansen et al., 2009; McKeefry et al., 2007; Ayama et al., 2004; Sakurai et al., 2003). Last, although it is unclear if it is affected by the chromaticity of the stimuli, susceptibility to crowding is possibly the clearest behavioral deficit of peripheral vision (Rosenholtz, 2016). We note though that changes in behavioral performance do not always unidirectionally decrease with eccentricity; for example, the central fovea is tritanopic because of a lack of S-cones (Williams, MacLeod, & Hayhoe, 1981; see also Appendix D, Control Experiment 1), and textural segregation performance peaks outside of the fovea (Yeshurun & Carrasco, 1998, 2000).

Presumably, it is these psychophysical facts that motivate the claims that peripheral color perceptual experience is illusory. Before critiquing these, it is worth noting that there are at least two possible interpretations of these statements, hinging on what is meant by the term "illusory": *(1) Perceptual performance for detection, discrimination, and characterization of colored stimuli in the periphery is dissociated from expected performance, that is, we are metacognitively overconfident about our peripheral color perception performance; or (2) our knowledge of our subjective experience of peripheral color at each moment is fragile and easily revised or distorted by expectation and memory. This happens to such an extent that the subjective reports of naive participants may completely dissociate from their actual color experience*. We are unsure of whether the skeptics mean to defend either or both of these claims, but as we believe both are inaccurate, we shall address each in turn.

On the first interpretation: Variation in perceptual abilities across the visual field does not imply an outright lack of perception per se. The lack of high-acuity chromatic perception in the periphery does not suggest a complete absence of chromatic perception, no more than an inability to detect microscopic stimuli presented foveally implies an absence of foveal vision (Anstis, 1998). Similarly, formal modeling of variation in chromatic perception across the visual field suggests that natural stimuli should still appear colorful at peripheral eccentricities (Haun, 2021). Many previous studies have shown that chromatically presented stimuli can be detected when presented peripherally, and even identified with foveal stimuli if appropriately magnified (Hansen et al., 2009; Abramov et al., 1991). Our current study supports this previous work and adds to it the finding that the perception of

the relational properties of these stimuli is equivalent when they are peripherally magnified.

Regarding the second: When considering the subjective experience of color as opposed to just stimulus detection, it is not clear that any experiment based simply on behavioral performance can be used to argue for an absence of peripheral color experience. For example, color constancy and metamers demonstrate the potential dissociation between stimulus properties and subjective experience (Gegenfurtner, Bloj, & Toscani, 2015; Adelson, 2000; Nimeroff & Yurow, 1965). In addition, people sometimes report color experiences in the complete absence of stimuli, such as in afterimages (van Boxtel, Tsuchiya, & Koch, 2010), the McCollough effect (Humphrey & Goodale, 1998), or even while dreaming (Kahn, Dement, Fisher, & Barmack, 1962). Thus, claims of an absence of color experience in peripheral vision based on failures of stimuli detection could just as easily be used to deny the presence of color experiences, despite subjective reports to the contrary, in all these other circumstances.

Given all this, what then explains the seeming discrepancy between our findings and the failure of participants to notice the complete absence of peripheral color stimuli in the virtual reality experiments of Cohen and colleagues (2020)? We believe the answer is a combination of how our differing paradigms exploited attention and expectation. Our experimental setup involved participants observing briefly presented colored stimuli in random locations against an otherwise featureless gray background, with participants discouraged from making saccades. This paradigm presumably enhances the salience of the peripheral stimuli through their sudden onset and lack of predictable location, engaging exogenous attention. In contrast, Cohen's virtual reality display involved participants observing colorful natural scene imagery, with participants free to saccade, for multiple seconds before gradual peripheral desaturation. Notably, participants often failed to notice peripheral desaturation even when they were aware of its upcoming possibility, in a manner reminiscent of gradual change blindness (Simons, Franconeri, & Reimer, 2000). We suspect this is likely because of memory of the scene colors along with a strong prior perceptual expectation that naturalistic scenes have constant physical color. One possible result of these factors is that participant expectation of constant peripheral color results in subjectively experienced color even when stimulus color is removed, that is, "filling in" (Balas & Sinha, 2007; Komatsu, 2006). Another possibility is that participants consciously experience the desaturation but fail to report it because of a failure to attend to and subsequently remember their peripheral experiences, that is, inattentional blindness or amnesia (Chen & Wyble, 2016; Mack, 2003; Wolfe, 1999). Either way, although we agree with Cohen's assertion that the participants can have surprisingly nonveridical color expectation, memory, and reports under these conditions, we disagree that this necessarily implies "the immediate impression of a rich, colorful experience that encompasses their entire visual world…is surprisingly inaccurate." (Cohen et al., 2020).

Nonetheless, both virtual reality headsets and our own Panoworks display provide a means to perform far-field psychophysics with significant advantages over historical approaches. Both approaches allow for easy manipulation of stimulus duration, size, and type in combination with head fixation and eye tracking. One interesting question that remains using both techniques is what is the source of individual differences in subjective reports across the different stimulus parameter conditions (Figure A3). Potential candidates include differences in: cortical magnification (Freeman & Simoncelli, 2011); color sensitivity in the periphery, perhaps because of interindividual cone receptor density variations (Afraz et al., 2010), or attention.

Last, we wish to conclude with the connections between our experimental approach and recent developments in the philosophy of consciousness. A number of philosophers have suggested the conceptual possibility of characterizing phenomenal consciousness through obtaining the structural relationships between experiences (Lee, 2021; Rosenthal, 2015; Chalmers, 1996; Nagel, 1974), with some going so far as to say these relations provide a complete description of any experience (Lyre, 2022; Fink et al., 2021). A similar claim of the feasibility of using a relational approach to characterize qualia has been made from a more formal category theoretical perspective (Tsuchiya & Saigo, 2021; Tsuchiya, Taguchi, & Saigo, 2016). These theoretical research programs go hand in hand with the popular neuroscientific method of representational similarity analysis (Shinkuma, Nishida, Kado, Maeda, & Nishimoto, 2019; Kriegeskorte & Kievit, 2013), among other efforts to pursue neuroscience of consciousness research related to qualia spaces (Lau, Michel, LeDoux, & Fleming, 2022; Tallon-Baudry, 2022). Our experimental approach provides a method for exploring these ideas empirically in a way that has actual consequences for these relational theories. Future studies that would also have interesting philosophical implications while making use of this approach could extend this work into other modalities than vision, characterize nonneurotypical populations such as those with color blindness, or collect subjective relationship reports in combination with neural recordings.

APPENDIX A

Table A1. Hypothesis Summary Table

Question	*Hypothesis*	*Sampling Plan*	*Analysis Plan*	*Interpretation Given to Different Outcomes*
1. Are the overall structures of color dissimilarity reports at different eccentricities the same?	H1A) Participants presented with peripherally magnified physically equivalent stimuli will provide CC & PP responses that are positively correlated.	Recruitment will stop when any of the following are met: • The BF for the hypothesis exceeds 10 • The BF for hypothesis is below 0.1 • 50 participants have been recruited. Participants will be excluded if: they self-report visual impairments (including glasses); fail to complete all trials; or score below 77% on catch trials. All other participants will be included.	We will calculate the Pearson correlation of CC and CP judgments for each participant. We will first Fisher-*Z* transform these correlations and then test whether the mean of their distribution is nonzero by performing a Bayesian one-sample *t* test using the default Jeffreys prior for the variance and Cauchy prior for the effect size.	• The results are consistent with the hypothesis if the BF for the model intercept is ≥ 3. • The results are consistent with the null hypothesis if the BF for the model intercept is ≤ 1/3. • The results will be taken to be inconclusive if 1/3 ≤ BF ≤ 3
	H1B) As above, but for CC & CP	As above	As above, but for CC & CP	As above
	H1C) As above, but for CP & PP	As above	As above, but for CP & PP	As above
1. Are physically equivalent stimuli presented at differing eccentricities subjectively equivalent?	H2A) Participants presented with peripherally magnified physically equivalent stimuli will provide dissimilarity responses closer to zero than expected from random responding in the CC condition.	As above	The difference in dissimilarity responses to identically vs. non-identically colored stimuli will be converted to a normalized trace difference for each participant in the peripherally magnified groups. We will fit this data from the CC condition to a Bayesian beta regression model with no regressors to calculate the intercept. We will use the default prior of a mean of 0 and a standard deviation of 2.5, which corresponds to a prior belief that participants will respond equivalently for identical and non-identical stimuli.	As above
As above	H2B) As above, but for the CP condition	As above	As above, but for the CP condition	As above
As above	H2C) As above, but for the PP condition	As above	As above, but for the PP condition	As above

2. Does failure to peripherally magnify small stimuli prevent color equivalence in the periphery?	H2D) Participants presented with small, unmagnified stimuli in the periphery that are physically equivalent will provide dissimilarity responses larger than those of participants shown larger or peripherally magnified stimuli, i.e., that there will be an interaction effect between size and magnification.	As above	The difference in dissimilarity responses to identically vs. non-identically colored stimuli will be converted to a normalized trace difference for each participant. We will fit the subset of this data corresponding to PP judgments with a Bayesian beta regression model with size, magnification and Size × Magnification as regressors to calculate the coefficient for the Size × Magnification regressor. We will use the default prior that corresponds to a prior belief that the coefficient for all of the regressors is zero.	As above
3. Are color-pairs subjectively equivalent regardless of presentation eccentricity?	H3) Participants will provide the same dissimilarity responses to the same pair of colored stimuli under CC, CP, or PP conditions, and hence the mean variance in responses across the conditions across all color pairs will be closer to zero than expected from random responding.	As above	For each participant in the peripherally magnified groups, variance in the dissimilarity responses to pairs of colored stimuli across the CC, CP, and PP conditions will be calculated, and the mean of this variance across all color pairs obtained. We will also calculate a participant's mean variance in the case where the color pair labels for the judgments are scrambled for the CP and PP eccentricities. We will then take the difference between the scrambled and original mean variance case for each participant as a measure of their judgment consistency across eccentricities. We will fit this mean variance difference data to a Bayesian beta regression model with no regressors to calculate the intercept. We will use the default prior of a mean of 0 and a standard deviation of 2.5, which corresponds to the prior belief that randomizing the structure of responses in different eccentricities will not affect the consistency of participant judgments across eccentricities.	As above

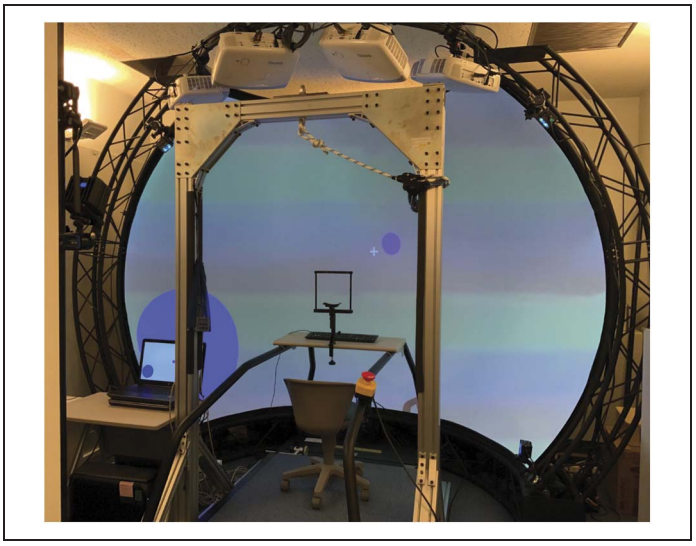

Figure A1. Panoworks display and setup. Horizontal streaking is an artifact of the camera and not perceived by participants.

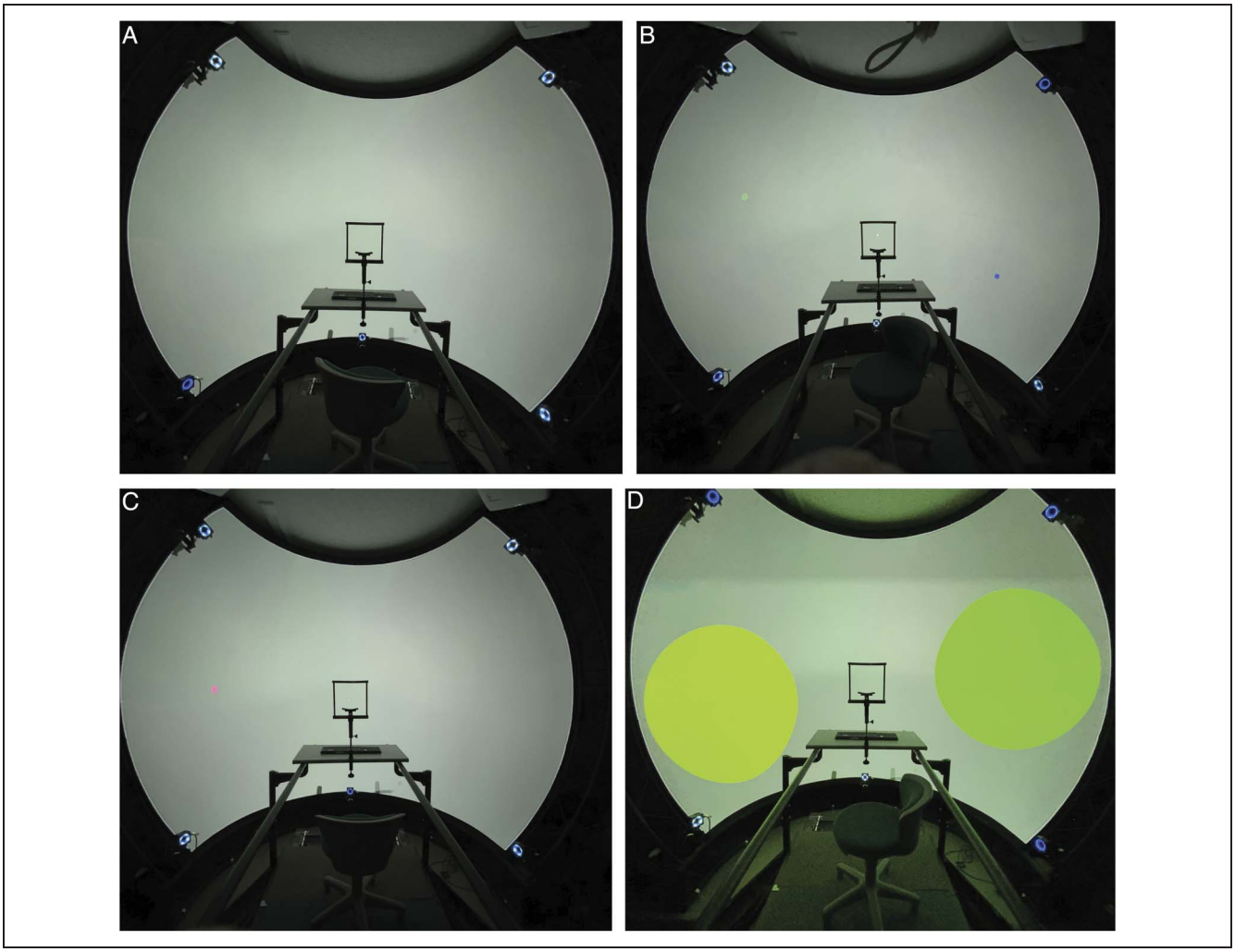

Figure A2. Panoworks display and setup for each condition. (A) FIX SMALL during a PP trial presentation. Stimuli have a 0.1 DVA and are presented at 60 DVA. The stimuli are too small to be easily seen in this image. (B) FIX BIG during a PP trial presentation. Stimuli are again presented centered at 60 DVA but with a diameter of 2 DVA. (C) MAG SMALL during a CP trial. The central stimulus is presented centered at 1 DVA with a diameter of 0.1 DVA and is too small to see on this image. The peripheral stimulus is centered at 60 DVA with a 2.7-DVA diameter. (D) MAG BIG during a PP trial. Both stimuli are presented centered at 60 DVA with a 54-DVA diameter.

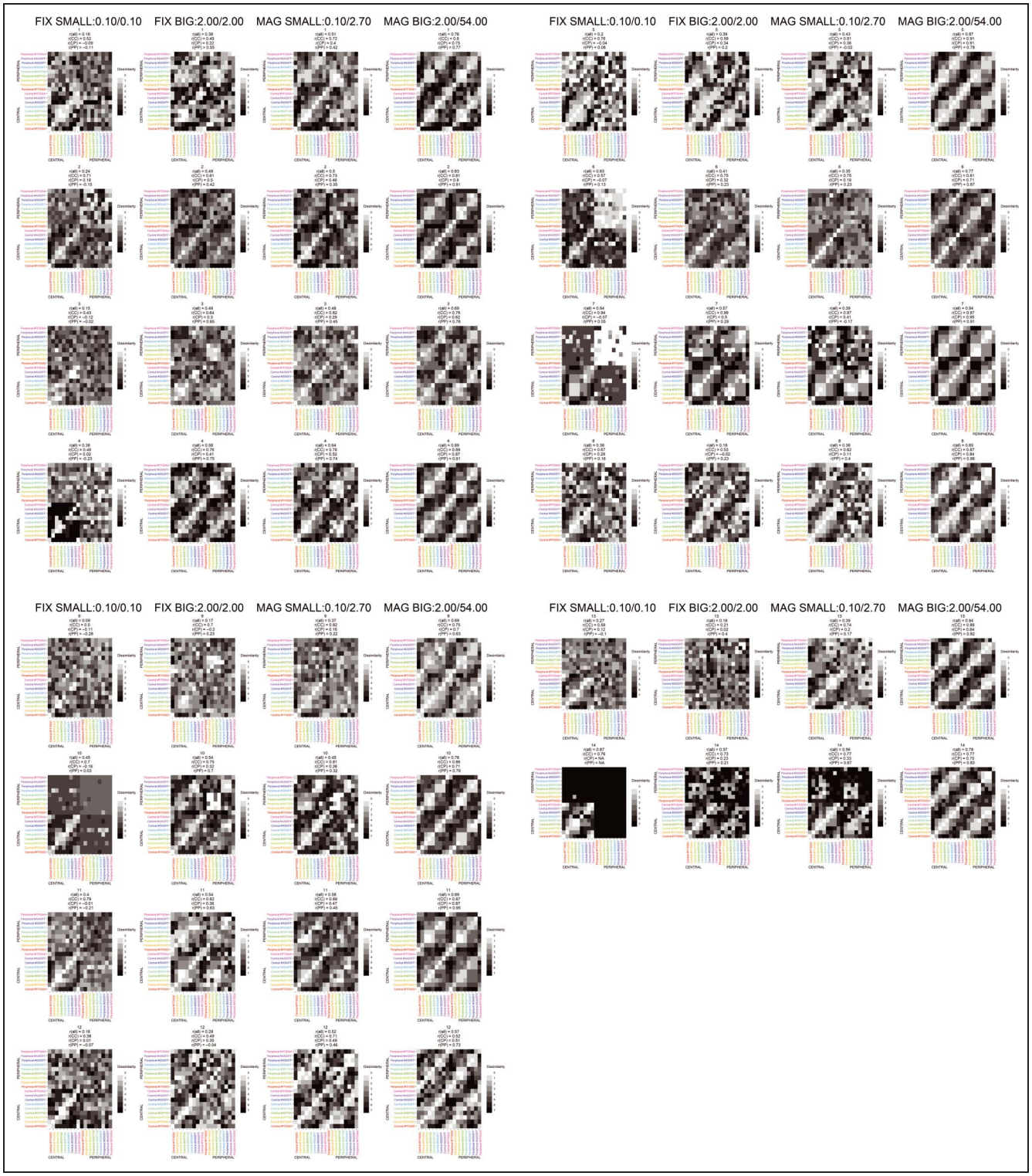

Figure A3. Individual raw dissimilarity values from the in-laboratory experiment placed into matrices for each participant. Each row is an individual participant, and each column corresponds to a stimulus parameter condition. The plots are asymmetric as participants were presented with each (non-identical) color pair twice and each response is plotted. The numbers above each plot show the double-pass correlation for each eccentricity condition for each stimulus parameter condition per participant.

APPENDIX B

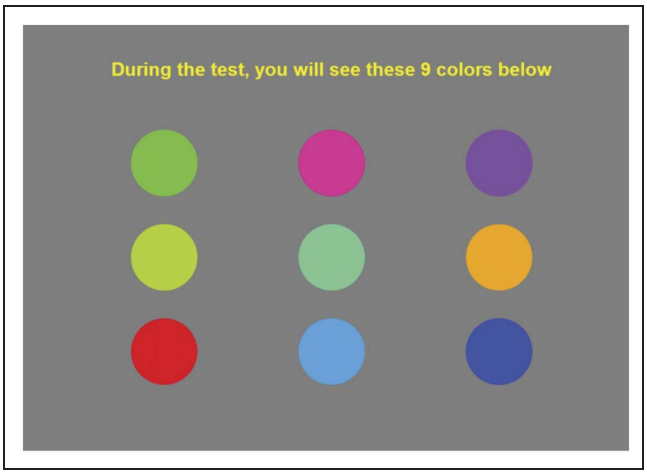

Figure B1. A sample color grid presented at the start of the experiment to familiarize participants with the colors they will be comparing. The colors were presented in random locations in the grid for each participant.

Table B1. Stimuli Used throughout the Experiments

Hex Code	*HSV*	*Name*
#7f7f7f	[0, 0, 0.5]	Background
#ff0000	[0, 1, 1]	Color 1
#ffaa00	[40, 1, 1]	Color 2
#aaff00	[80, 1, 1]	Color 3
#00ff00	[120, 1, 1]	Color 4
#00ffa9	[160, 1, 1]	Color 5
#00a9ff	[200, 1, 1]	Color 6
#0000ff	[240, 1, 1]	Color 7
#aa00ff	[280, 1, 1]	Color 8
#ff00aa	[320, 1, 1]	Color 9

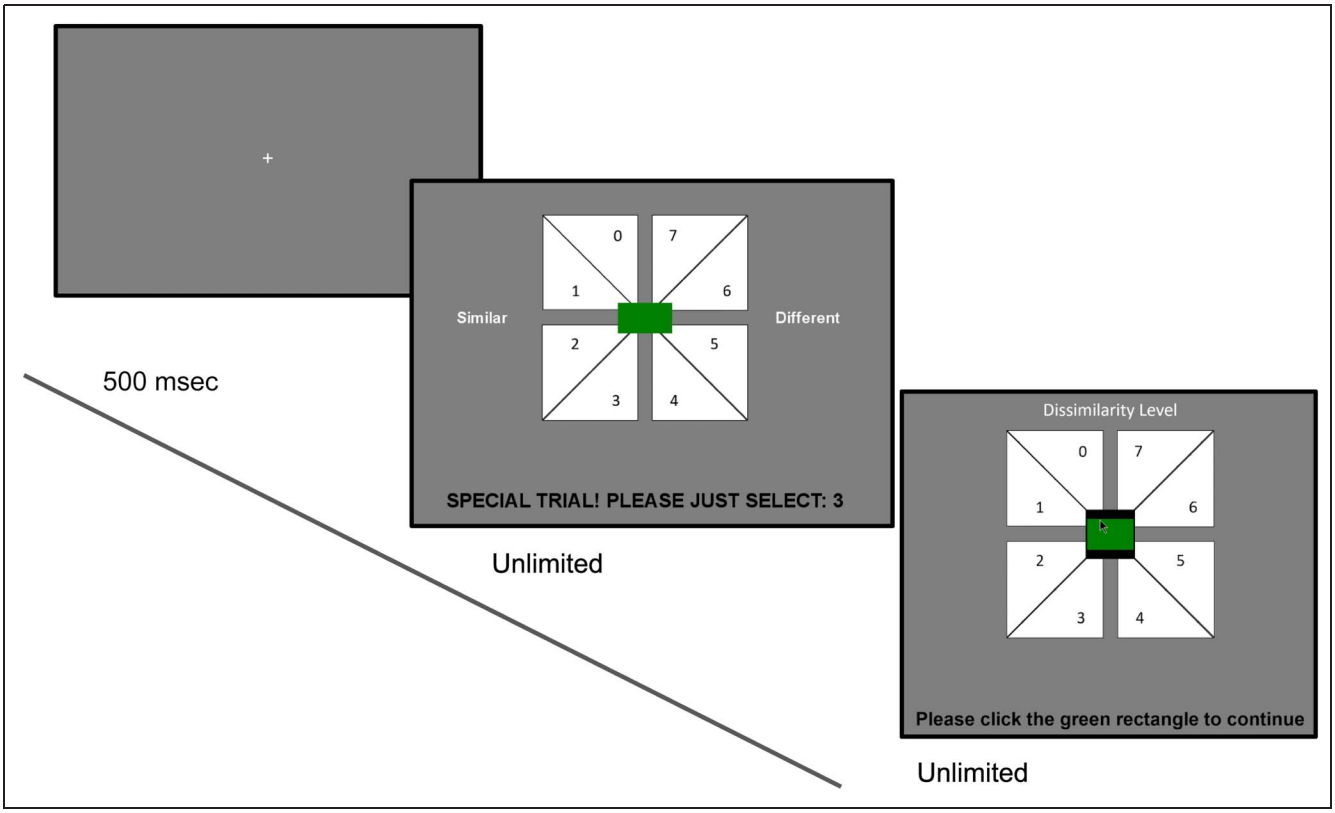

Figure B2. Task procedure for catch trials. Ten catch trials are randomly inserted into the main trial sequence. After the 500-msec fixation, instead of stimuli display, participants are instead shown a response screen where they are directed to select a particular value chosen at random and varying between catch trials. Participants are then directed to click on the green rectangle at the center of the screen to proceed to the next trial.

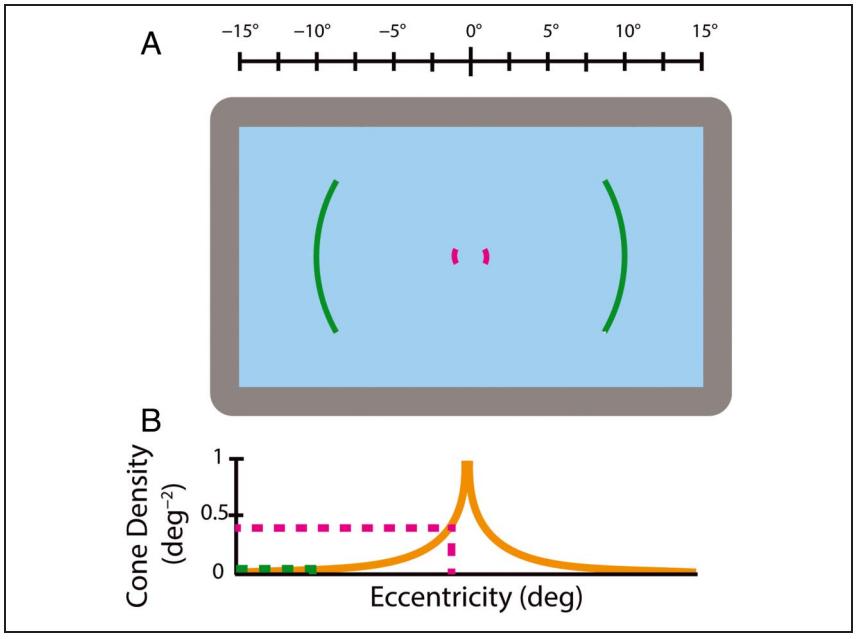

Figure B3. (A) Schematic for the pilot online experiments. The panel shows the achievable eccentricities of stimuli with typical computer monitors. Central and peripheral stimuli are presented at random locations ±30° of horizontal and centered on the pink (10 DVA) and green (1 DVA) lines, respectively. (B) Cone photoreceptor densities at central and peripheral eccentricities relative to the retinal maximum (Watson, 2014; Curcio et al., 1990). Eccentricities for (A) and (B) are shown by the scale above (A).

APPENDIX C

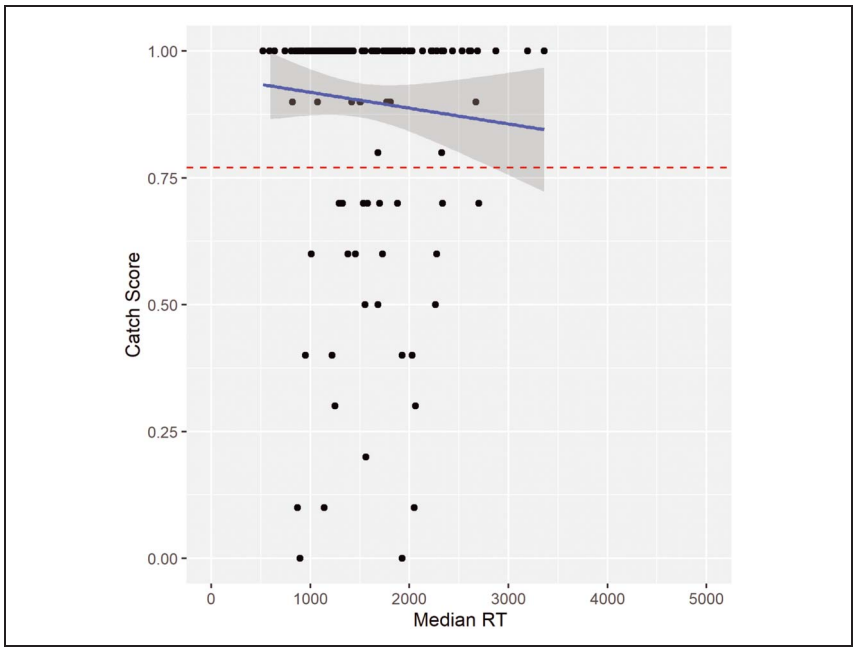

Figure C1. Pilot online results. Catch trial accuracy (chance performance 12.5%) plotted against the mean RT on main trials for a participant. The horizontal red line indicates the catch score cutoff of > 0.77. Only participants above the cutoff were included in the pilot analysis.

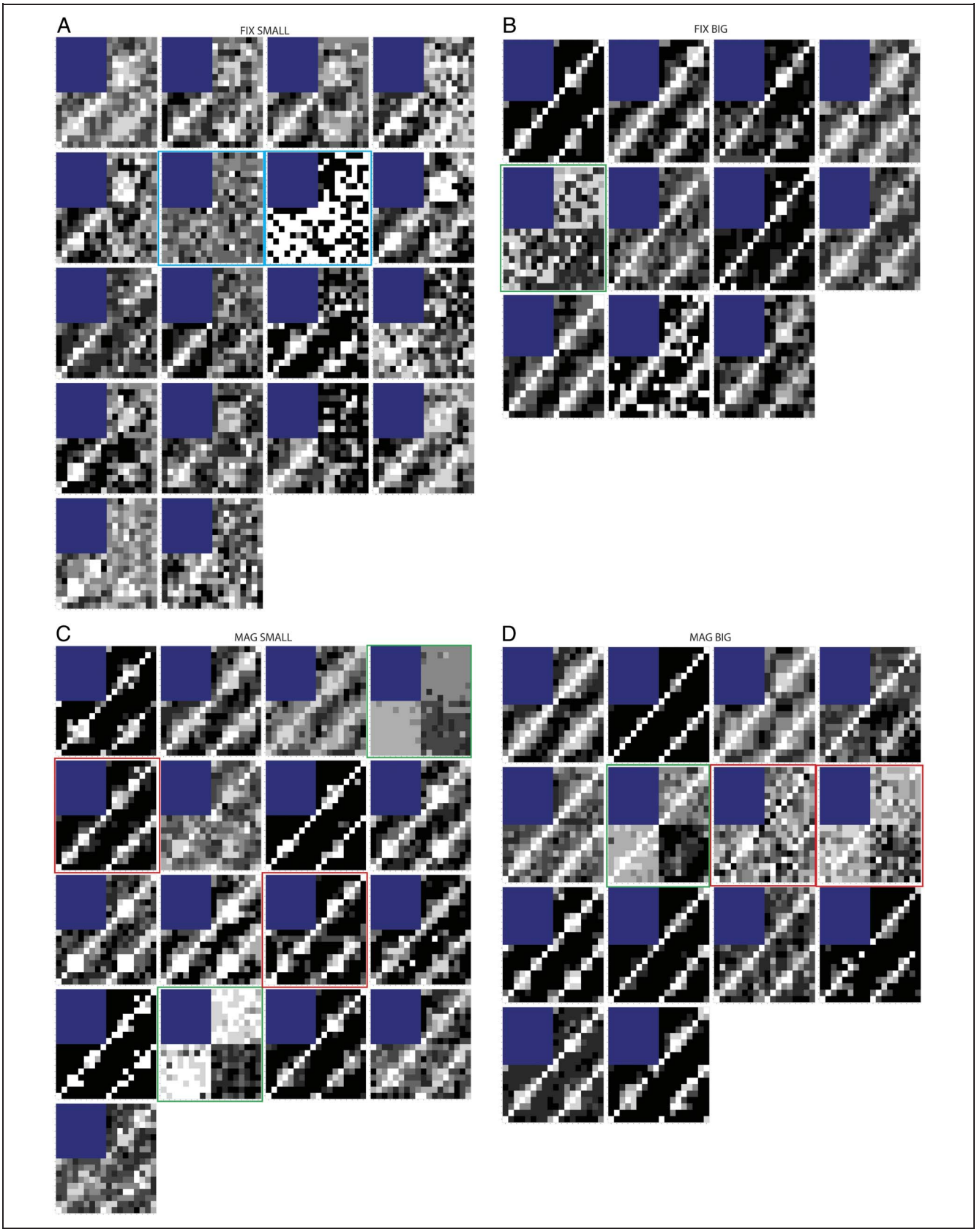

Figure C2. Individual raw dissimilarity values from the online pilot experiment placed into matrices for each participant in the (A) small, fixed-size stimuli; (B) large, fixed-size stimuli; (C) small, peripherally magnified stimuli; and (D) large, peripherally magnified stimuli groups. Participants who appear to have misunderstood the response instructions and inverted their dissimilarity responses are marked in red. Participants who otherwise have CC normalized traces over 3.5 are marked in blue. Participants marked in red or blue had their responses inverted for the pilot analysis. Participants who otherwise are suspected of misunderstanding the instructions and/or task noncompliance are marked in green.

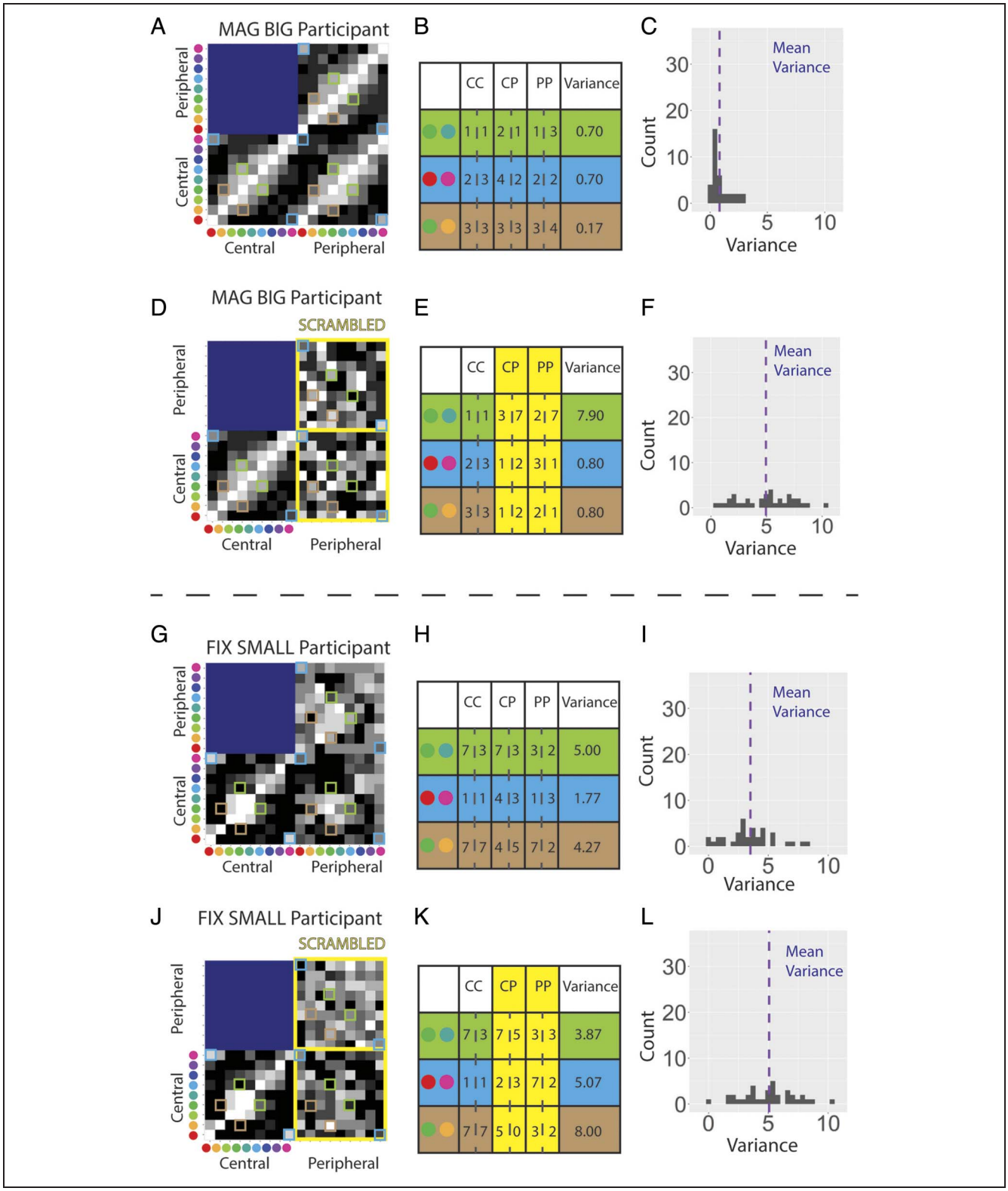

Figure C3. A demonstration of how mean variance is calculated for each participant. (A) The raw dissimilarity values for a representative participant from the MAG BIG experimental condition. (B) An illustration of how the variance in dissimilarity responses is calculated for three different color pairs. The background color of each row marks the six dissimilarity values marked by the corresponding colored boxes in (A). (C) A histogram of the variance for all color pairs for the participant. (D) The same participant as in (A), but with their similarity judgments for the CP and PP conditions randomly scrambled. (E) The same variance calculations as per (B) but with the scrambled values for CP and PP. (F) The variance in similarity judgments for all color pairs in the given CP and PP shuffling. The mean variance of the distribution is considerably **higher** than in (C), indicating the original similarity judgments were **consistent** across eccentricity conditions. (G–L) as per (A–F), but for a participant in the FIX SMALL group. Note that the mean variance is **similar** in both the original (I) and scrambled (L) cases, indicating that the participant's similarity judgments were **inconsistent** across eccentricity conditions.

Table C1. Normalized Trace Difference per Condition from the Online Pilot Experiment (n = 31) Using Peripherally Magnified Stimuli Shown at 10 DVA

Condition	BF_{10}
CC, CP	$> 10^{14}$
CC, PP	$> 10^{13}$
CP, PP	$> 10^{13}$

Table C2. Normalized Trace Difference per Condition from the Online Pilot Experiment (n = 31) Using Peripherally Magnified Stimuli Shown at 10 DVA

Condition	*Mean*	*95% HPD*	BF_{10}
CC	3.31	2.49 to 3.99	$> 10^3$
CP	2.93	2.10 to 3.65	$> 10^3$
PP	3.40	2.60 to 4.07	$> 10^3$

Table C3. Normalized Trace Difference Data from All Participant Groups in the Online Pilot Experiment (n = 31) PP Condition Fitted to a Bayesian Beta Regression

	β	*95% HPD*	BF_{10}
Intercept	−0.1	−0.47 to 0.27	
Size	1.2	0.57 to 1.86	
Magnification	0.3	0.15 to 0.39	
Size × Magnification	−0.2	−0.44 to −0.05	1.49

Table C4. Mean Variance from the Online Pilot Experiment (n = 31) Using Peripherally Magnified Stimuli Shown at 10 DVA

Mean	*95% HPD*	BF_{10}
2.65	1.27 to 3.96	20.44

APPENDIX D

Below, we describe methods related to the online pilot experiments.

Pilot Experiment: Online Version of the Main Experiment

Ethics

Experimental procedures were approved by the Monash University Human Research Ethics Committee (Project ID: 17674). Participants were provided electronically with written consent forms before the commencement of the experiment and provided electronic consent to participate. Participants were compensated for their time at a rate of $6/hour.

Design

Participants. Participants were recruited remotely through CloudResearch (Litman, Robinson, & Abberbock, 2017), a virtual wrapper for Amazon's Mechanical Turk (MTurk) platform. Participants accessed the experiment through the MTurk worker area and provided data using their own personal computers. Only English native speakers from Australia, Canada, India, and the United States of America were recruited. Participants were familiar with the online MTurk platform as they were only invited to participate if they had performed at least 1000 unrelated tasks online with at least 97% approval. Participants were only recruited from those with a Windows operating system.

Display apparatus. Because of the nature of online experimentation, participants used their own computer screen to perform the experiment. Screen properties necessary for appropriate stimulus display were obtained via calibration steps at the beginning of the task, described below in Procedure section. Stimuli were presented via Inquisit 6 (millisecond) display software. Centrally presented stimuli were presented at 1 DVA as per the main methods. Peripherally presented stimuli in the online experiment were centered at 10 DVA, outside of the macula's visual field and thus in the near periphery (Figure B3).

Consent. Once participants agreed to participate in our study in the TurkPrime interface, they were provided with a link that directs them to our study hosted by Millisecond. The first page of the experiment was a consent form that they could electronically sign by pressing the spacebar. Participants were informed that the data collection process was anonymous and that they could quit the experiment at any time.

Screen size calibrations. Following the consent page, participants were directed through a series of steps to obtain both their physical screen size and viewing distance from the screen and were instructed that they must perform these properly to complete the task. The following protocol has been adapted from a previous study validating this method (Li, Joo, Yeatman, & Reinecke, 2020). Participants were asked to place a card (e.g., credit card, student card, or any other card that is the same dimension as a credit card: 8.6 cm × 5.4 cm) against the screen and adjust a rectangle on their screen to match the size of the physical card. As these cards come in a standard size, this allowed us to infer the screen size from the number of pixels in the displayed rectangle. Participants were then directed to fixate on a point on the right-hand side of their

screen with their left eye while closing their right eye. Participants were then shown a small circle that moves from right to left and asked to press the ENTER key when the circle disappears from view, which occurs upon entering the participant's blind spot. As the blind spot is located about 13.5° from the center of vision, this allowed for trigonometric calculation of the viewing distance if the screen size is known. If the calibration values suggest the participant would be unable to complete the experiment because of inadequate screen size, participants are rejected from the experiment. For a demonstration of the calibration procedure, see the online version of the task: https://mili2nd.co/frub.

Task. Online participants performed trials with only one stimulus size and peripheral magnification setting. Each participant completed 81 CC, 81 CP, and 81 PP main trials and 10 catch trials, equaling 253 trials. All other methodological details are as per the main text.

Sampling Plan

Participant exclusion. Participant quality control occurred both before and after the main experiment. Before the main experiment and after the calibration procedure, we stopped the experiment if participants' estimated screen size did not allow us to present the peripheral stimuli without being cut off by the edge of the screen. In addition, we excluded participants whose estimated screen diameter is less than 26 cm or viewing distance is less than 30 cm, as these values would suggest the task is being performed on a phone or tablet.

Following the main experiment, we calculated the mean and standard deviation of the participant catch scores. We excluded participants who scored less than 1 *SD* below the median (77%) on catch trials, as this is suggestive of inattentive behavior (Figure C1). Participants who did not complete all of the main experimental trials were also excluded. All other participants were included in the data analysis.

Analysis Plan

Preprocessing. Upon visual inspection of the online data, it appeared that some participants provided dissimilarity responses that were inverted relative to what would be expected even in central vision presentation conditions, presumably because of misunderstanding the response method. As a result, for the pilot data, we inverted the dissimilarity responses of participants with a CC normalized trace > 3.5. We will not perform this for the laboratory-based experiment should it occur then. Instead, we have optimized the task instructions and added practice trial feedback to make participant misunderstandings less likely. All other analyses are the same as per the main text.

Control Experiment 1: Similarity Judgments in the Central Fovea

As a check of the efficacy and robustness of our methods, we performed a control experiment to assess the impact of the absence of blue-sensitive S-cones in the central retina on color similarity judgments. The central fovea is similar to the condition of tritanopia, which results in impairments in blue-green color discrimination (Williams et al., 1981). We attempted to replicate these previous findings using our similarity judgment technique.

Online participants performed a modified version of our color similarity judgment task, where stimuli varied either from blue to green (test) or blue to red (control). Before commencing the task, participants performed a heterochromatic flicker adjustment task to equilibrate the luminance values for their specific monitor (Wagner & Boynton, 1972; Figure D1A). This was intended to increase the likelihood that stimuli varied subjectively in chroma rather than in luminance. Participants also performed the screen size calibrations as per Experiment 1. Following the calibrations, participants performed a modified version of our original task (Figure D1B). Stimuli were presented under two conditions: 0°/1° or 1°/1°. At the 0° position, stimuli with a diameter of 0.1° were presented centrally at 0°, which is free of S-cones (Magnussen et al., 2004). In the 1° position, stimuli with a diameter of 0.45° are presented in a region of the fovea with S-cones. The fixation cross was changed to a fixation circle to avoid overlap with the 0° stimuli. The stimuli presentation time was changed from 250 msec to 100 msec to help prevent microsaccades. All other aspects, including the response protocol, remained the same as the previously reported experiments. First, we the authors performed the task, which was also replicated by $n = 14$ online participants with similar results.

Figure D1C (top) shows the dissimilarity rating matrix as expected from tritanopic, blue-green blind vision. All pairs of isoluminant blue-green patches between 0° and 1° patches look similar (i.e., the mean rating of ~2 out of 7), with an indistinct relationship structure. In contrast, participants gave a much clearer and structured dissimilarity rating matrix as expected from normal vision when they compared two small color patches both at 1° (Figure D1C, bottom). Participant dissimilarity responses to physically identical pairs (the mean of the diagonal, i.e., normalized trace) were much smaller when both stimuli were displayed on retinal areas with S-cones present (Figure D1; $p = .014$). A control experiment with isoluminant red-green patches (Figure D1E and F) confirmed that this pattern was not observed ($p = .985$).

Control Experiment 2: Examining the Possibility of Screen Chromatic Heterogeneity

We examined the effects of possible variation in the chromatic properties on a participant's monitor. To do so, we

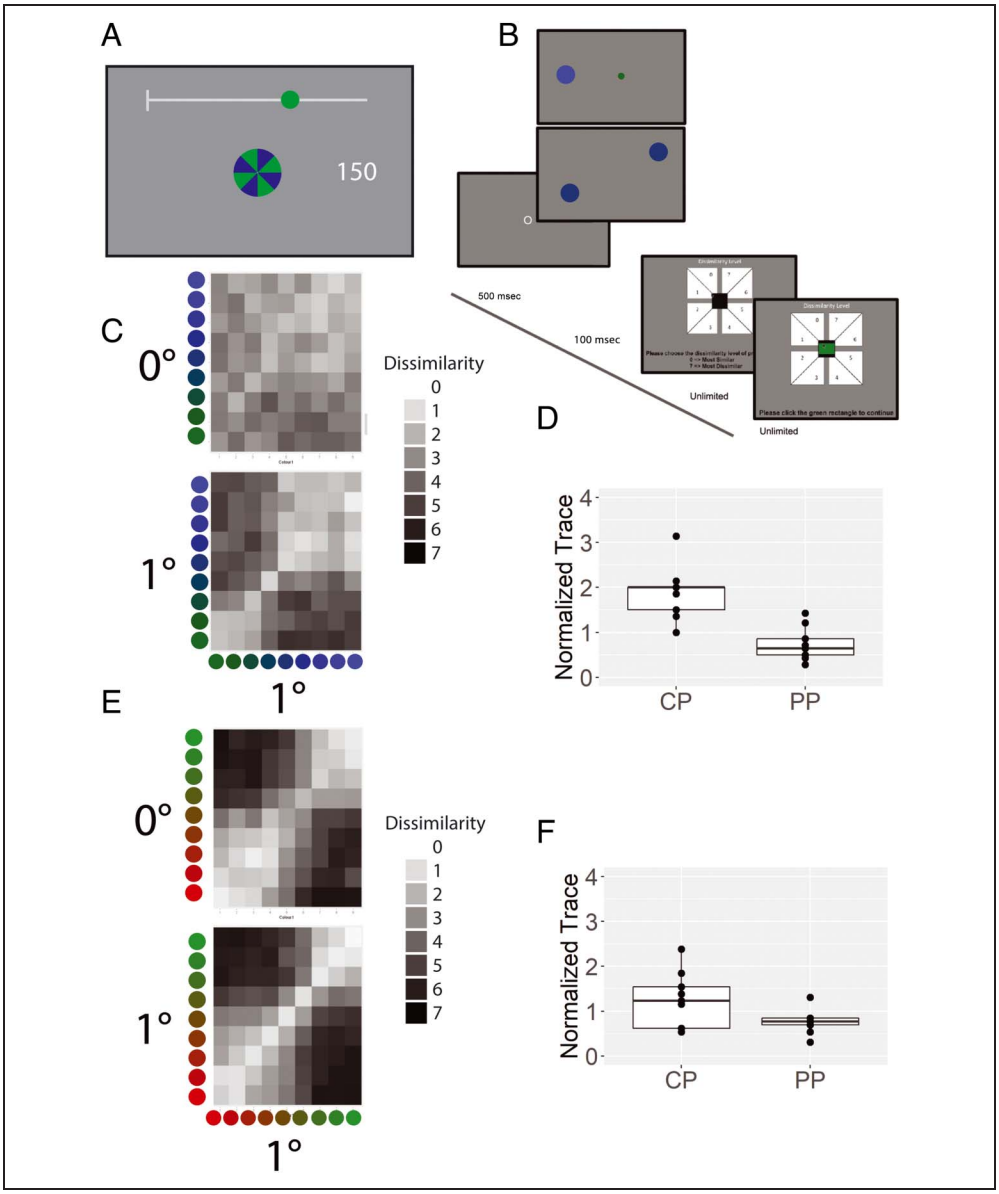

Figure D1. Color similarity judgments in the central fovea. (A) Participant-specific color equiluminance values were obtained via the minimal-flicker technique. Participants were shown a stimulus flickering between two solid colors (either blue/green or red/green) in the center of the screen and asked to adjust the brightness of one of the colors using a slider until the flickering was minimized and the colors appeared equiluminant. (B) The main similarity judgment task followed the luminance and screen size calibrations. The calibrations allowed participants to be shown a set of stimuli that were isoluminant on their monitors at the center of the screen. Stimuli presentation was preceded by a fixation circle (500 msec) with an interior larger than the central stimulus. (C) Group-mean dissimilarity judgment matrices for experiences generated by equiluminant stimuli varying from green to blue in 0° versus 1° comparison (top) and 1° versus 1° (bottom) conditions. Note that stimuli were always displayed within the fovea. (D) The normalized traces from the dissimilarity matrices for participants show stimuli that varied from blue to green. Participants rated the physically identical blue–green pairs as more different (i.e., rating of ~2) in the 0° versus 1° comparison condition but more similar between the 1° and 1° comparison condition (i.e., ~0.7). (E and F) The same as (C) and (D) but for red to green. Red–green pairs looked similar at both the 1° versus 1° (i.e., ~0.7) and 0° versus 1° (i.e., ~1.2). Boxplots centerline is median, box shows 25th–75th percentile, and whiskers show the largest value within 150% of the interquartile range. n = 13–14.

modified Experiment 1, systematically varying the retinal or screen location of a stimulus (Figure D2). Online participants performed color similarity judgments for stimuli presented under 2 (screen location: fixed or not) × 2 (retinal location: fixed or not; Figure D2A and B). Each participant provided a full set of judgments for all color pairs under each of the four conditions, with the trials for different conditions randomly interspersed. Three different stimulus presentation points were selected for the experiment, each 5° apart. Each trial had four phases. First, a fixation cross was presented at the left side of the screen for 500 msec, followed by a stimulus for 250 msec in one of two possible locations. After this, a fixation cross (500 msec) was presented in one of two locations, followed by an additional stimulus (250 msec) in one of two possible locations depending on the second fixation point. This design meant that stimuli always had a 50% probability of presentation to the left or right of a fixation cross after presentation.

The same pattern of similarity relationships are observed regardless of condition (Figure D2C). This suggests that chromatic inhomogeneity of commercial displays does not affect the pattern of similarity judgments provided. We note also that as our methods are based on analyzing the structure of sets of similarity judgments (e.g., rather than binary classification of stimuli), they are robust to minor changes in stimulus properties that lead to only minor changes in similarity structure.

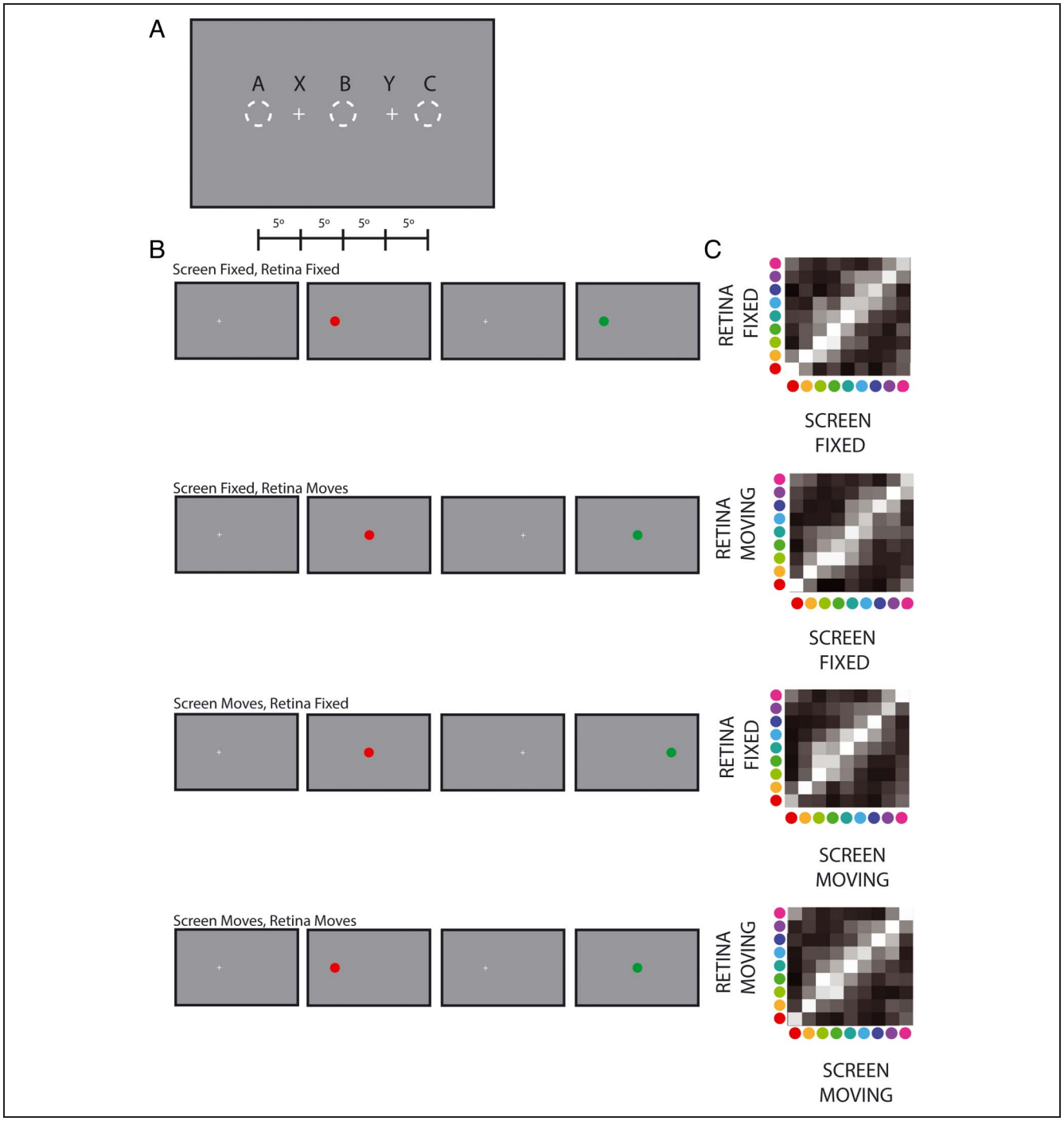

Figure D2. Examining the possibility of screen chromatic heterogeneity. Online participants performed color similarity judgments for stimuli presented under four different conditions. Each participant provided a full set of judgments for all color pairs under all four conditions, with the trials for different conditions randomly interspersed. (A) Two different fixation points and three different stimulus presentation points were selected for the experiment, each 5° apart. (B) Illustration of the four possible conditions for stimuli presentation: screen location fixed, retinal location fixed, neither, or both. (C) The mean dissimilarity judgments for each color pair across participants. Each participant performed all four conditions. The same pattern of similarity relationships are observed regardless of condition. $n = 7$.

Acknowledgments

A. Z.-J./N. T. are supported by Australian Research Council (DP180100396), National Health and Medical Research Council (APP1183280), and the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation (FQXi-RFP-CPW-2017). N. T./M. Y. are supported by Grant-in-Aid for Transformative Research Areas (B; N. T.: 20H05710, M. Y.: 20H05711) from Japan Society for the Promotion of Science. M. Y. is supported by Moonshot R&D Grant (JPMJMS2295-01) from Japan Science and Technology Agency and KAKENHI (22H01108, 22 K18265) from Japan Society for the Promotion of Science. The funders have/had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would additionally like to thank Ruitong Fan for helping develop the initial pilot experiment.

Reprint requests should be sent to Ariel Zeleznikow-Johnston, Turner Institute for Brain and Mental Health & School of Psychological Sciences, Faculty of Medicine, Nursing, and Health Sciences, Monash University, 770 Blackburn Rd., Melbourne, Victoria, Australia, 3800, or via e-mail: ariel.zeleznikow-johnston@monash.edu, or Makiko Yamada, Institute for Quantum Medical Science, National Institutes for Quantum and Radiological Science and Technology, 4-9-1 Anagawa, Inage-ku, Chiba, 263-8555, Japan, or via e-mail: yamada.makiko@qst.go.jp.

Data Availability Statement

We published a preregistered research plan on the Open Science Framework before data collection (https://osf.io/rg7kf/). Our experimental code, raw data, and analysis code are all available on the Open Science Framework (https://osf.io/5pfrg/).

Author Contributions

Ariel Zeleznikow-Johnston: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Visualization; Writing—Original draft; Writing—Review & editing. Yasunori Aizawa: Data curation; Investigation; Project administration; Writing—Review & editing. Makiko Yamada: Funding acquisition; Project administration; Supervision; Writing—Review & editing. Naotsugu Tsuchiya: Conceptualization; Formal analysis; Funding acquisition; Methodology; Project administration; Supervision; Writing—Review & editing.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

Abramov, I., Gordon, J., & Chan, H. (1991). Color appearance in the peripheral retina: Effects of stimulus size. *Journal of the Optical Society of America A*, *8*, 404–414. https://doi.org/10.1364/JOSAA.8.000404, PubMed: 2007915

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: MIT Press.

Afraz, A., Pashkam, M. V., & Cavanagh, P. (2010). Spatial heterogeneity in the perception of face and form attributes. *Current Biology*, *20*, 2112–2116. https://doi.org/10.1016/j.cub.2010.11.017, PubMed: 21109440

Anderson, S. J., Mullen, K. T., & Hess, R. F. (1991). Human peripheral spatial resolution for achromatic and chromatic stimuli: Limits imposed by optical and retinal factors. *Journal of Physiology*, *442*, 47–64. https://doi.org/10.1113/jphysiol.1991.sp018781, PubMed: 1798037

Anstis, S. (1998). Picturing peripheral acuity. *Perception*, *27*, 817–825. https://doi.org/10.1068/p270817, PubMed: 10209644

Ayama, M., Sakurai, M., Carlander, O., Derefeldt, G., & Eriksson, L. (2004). Color appearance in peripheral vision. In B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings of the Human Vision and Electronic Imaging IX* (p. 260). https://doi.org/10.1117/12.522240

Balas, B., & Sinha, P. (2007). "Filling-in" colour in natural scenes. *Visual Cognition*, *15*, 765–778. https://doi.org/10.1080/13506280701295453

Bonnardel, V., Beniwal, S., Dubey, N., Pande, M., Knoblauch, K., & Bimler, D. (2016). Perceptual color spacing derived from maximum likelihood multidimensional scaling. *Journal of the Optical Society of America A*, *33*, A30–A36. https://doi.org/10.1364/JOSAA.33.000A30, PubMed: 26974936

Bosten, J. M., Robinson, J. D., Jordan, G., & Mollon, J. D. (2005). Multidimensional scaling reveals a color dimension unique to 'color-deficient' observers. *Current Biology*, *15*, R950–R952. https://doi.org/10.1016/j.cub.2005.11.031, PubMed: 16332521

Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception & Psychophysics*, *43*, 494–507. https://doi.org/10.3758/BF03207885, PubMed: 3380640

Carroll, S., & Dennett, D. C. (2020). *Daniel Dennett on minds, patterns, and the scientific image – Sean Carroll* (No. 78). https://www.preposterousuniverse.com/podcast/2020/01/06/78-daniel-dennett-on-minds-patterns-and-the-scientific-image/

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.

Chater, N. (2018). *Mind is flat: The remarkable shallowness of the improvising brain*. Yale University Press. https://doi.org/10.12987/9780300240610

Chen, H., & Wyble, B. (2016). Attribute amnesia reflects a lack of memory consolidation for attended information. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 225–234. https://doi.org/10.1037/xhp0000133, PubMed: 26348066

Cohen, M. A., Botch, T. L., & Robertson, C. E. (2020). The limits of color awareness during active, real-world vision. *Proceedings of the National Academy of Sciences, U.S.A.*,

117, 13821–13827. https://doi.org/10.1073/pnas.1922294117, PubMed: 32513698
Cohen, M. A., & Rubenstein, J. (2020). How much color do we see in the blink of an eye? *Cognition*, *200*, 104268. https://doi.org/10.1016/j.cognition.2020.104268, PubMed: 32473406
Cribari-Neto, F., & Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software*, *34*, 1–24. https://doi.org/10.18637/jss.v034.i02
Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *Journal of Comparative Neurology*, *300*, 5–25. https://doi.org/10.1002/cne.903000103, PubMed: 2229487
Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology*, *292*, 497–523. https://doi.org/10.1002/cne.902920402, PubMed: 2324310
Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology*, *159*, 203–221. https://doi.org/10.1113/jphysiol.1961.sp006803, PubMed: 13883391
Decock, L., & Douven, I. (2013). Qualia compression. *Philosophy and Phenomenological Research*, *87*, 129–150. https://doi.org/10.1111/j.1933-1592.2011.00545.x
Ennis, R. J., & Zaidi, Q. (2019). Geometrical structure of perceptual color space: Mental representations and adaptation invariance. *Journal of Vision*, *19*, 1. https://doi.org/10.1167/19.12.1, PubMed: 31573606
Fink, S. B., Kob, L., & Lyre, H. (2021). A structural constraint on neural correlates of consciousness. *Philosophy and the Mind Sciences*, *2*, 7. https://doi.org/10.33735/phimisci.2021.79
Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*, 1195–1201. https://doi.org/10.1038/nn.2889, PubMed: 21841776
Gegenfurtner, K. R., Bloj, M., & Toscani, M. (2015). The many colours of 'the dress.' *Current Biology*, *25*, R543–R544. https://doi.org/10.1016/j.cub.2015.04.043, PubMed: 25981790
Giron, C., Lau, H., & Knotts, J. D. (2018). Are open interviews superior to button presses? A commentary on Haun et al. (2017). *Symposium on the Brains Blog* [Blog post]. https://philosophyofbrains.com/2018/04/13/symposium-on-haun-tononi-koch-and-tsuchiya-are-we-underestimating-the-richness-of-visual-experience.aspx
Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). Rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1. https://mc-stan.org/rstanarm
Gordon, J., & Abramov, I. (1977). Color vision in the peripheral retina. II. Hue and saturation. *Journal of the Optical Society of America*, *67*, 202–207. https://doi.org/10.1364/JOSA.67.000202, PubMed: 839300
Hansen, T., Pracejus, L., & Gegenfurtner, K. R. (2009). Color perception in the intermediate periphery of the visual field. *Journal of Vision*, *9*, 26. https://doi.org/10.1167/9.4.26, PubMed: 19757935
Haun, A. M. (2021). What is visible across the visual field? *Neuroscience of Consciousness*, *2021*, niab006. https://doi.org/10.1093/nc/niab006, PubMed: 34084558
Helm, C. E. (1964). Multidimensional ratio scaling analysis of perceived color relations. *Journal of the Optical Society of America*, *54*, 256–262. https://doi.org/10.1364/JOSA.54.000256, PubMed: 14123941
Humphrey, G. K., & Goodale, M. A. (1998). Probing unconscious visual processing with the McCollough effect. *Consciousness and Cognition*, *7*, 494–519. https://doi.org/10.1006/ccog.1998.0369, PubMed: 9787058
Izmailov, C. A., & Sokolov, E. N. (1991). Spherical model of color and brightness discrimination. *Psychological Science*, *2*, 249–260. https://doi.org/10.1111/j.1467-9280.1991.tb00143.x
Kahn, E., Dement, W., Fisher, C., & Barmack, J. E. (1962). Incidence of color in immediately recalled dreams. *Science*, *137*, 1054–1055. https://doi.org/10.1126/science.137.3535.1054, PubMed: 14453103
Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews Neuroscience*, *7*, 220–231. https://doi.org/10.1038/nrn1869, PubMed: 16495943
Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*, 401–412. https://doi.org/10.1016/j.tics.2013.06.007, PubMed: 23876494
Lau, H., Michel, M., LeDoux, J. E., & Fleming, S. M. (2022). The mnemonic basis of subjective experience. *Nature Reviews Psychology*, *1*, 479–488. https://doi.org/10.1038/s44159-022-00068-6
Lee, A. Y. (2021). Modeling mental qualities. *Philosophical Review*, *130*, 263–298. https://doi.org/10.1215/00318108-8809919
Lennie, P. (1998). Single units and visual cortical organization. *Perception*, *27*, 889–935. https://doi.org/10.1068/p270889, PubMed: 10209632
Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a Virtual Chinrest. *Scientific Reports*, *10*, 904. https://doi.org/10.1038/s41598-019-57204-1, PubMed: 31969579
Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*, 433–442. https://doi.org/10.3758/s13428-016-0727-z, PubMed: 27071389
Logvinenko, A. D., & Maloney, L. T. (2006). The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics*, *68*, 76–83. https://doi.org/10.3758/BF03193657, PubMed: 16617831
Lyre, H. (2022). Neurophenomenal structuralism. A philosophical agenda for a structuralist neuroscience of consciousness. *Neuroscience of Consciousness*, *2022*, niac012. https://doi.org/10.1093/nc/niac012, PubMed: 36004320
Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current Directions in Psychological Science*, *12*, 180–184. https://doi.org/10.1111/1467-8721.01256
Magnussen, S., Spillmann, L., Stürzel, F., & Werner, J. S. (2004). Unveiling the foveal blue scotoma through an afterimage. *Vision Research*, *44*, 377–383. https://doi.org/10.1016/j.visres.2003.09.023, PubMed: 14659964
Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). BayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*, 1541. https://doi.org/10.21105/joss.01541
McKeefry, D. J., Murray, I. J., & Parry, N. R. A. (2007). Perceived shifts in saturation and hue of chromatic stimuli in the near peripheral retina. *Journal of the Optical Society of America A*, *24*, 3168–3179. https://doi.org/10.1364/JOSAA.24.003168, PubMed: 17912307
Mullen, K. T. (1991). Colour vision as a post-receptoral specialization of the central visual field. *Vision Research*, *31*, 119–130. https://doi.org/10.1016/0042-6989(91)90079-K, PubMed: 2006545
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, *83*, 435–450. https://doi.org/10.2307/2183914
Nimeroff, I., & Yurow, J. A. (1965). Degree of metamerism. *Journal of the Optical Society of America*, *55*, 185–190. https://doi.org/10.1364/JOSA.55.000185

Rajananda, S., Peters, M. A. K., Lau, H., & Odegaard, B. (2017). Subjective inflation of color saturation in the periphery under temporal overload. *bioRxiv*. https://doi.org/10.1101/227074
Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, *2*, 437–457. https://doi.org/10.1146/annurev-vision-082114-035733, PubMed: 28532349
Rosenthal, D. (2015). Quality spaces and sensory modalities. In P. Coates & S. Coleman (Eds.), *Phenomenal qualities* (pp. 33–65). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198712718.003.0002
Sakurai, M., Ayama, M., & Kumagai, T. (2003). Color appearance in the entire visual field: Color zone map based on the unique hue component. *Journal of the Optical Society of America A*, *20*, 1997–2009. https://doi.org/10.1364/JOSAA.20.001997, PubMed: 14620327
Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*, 128–142. https://doi.org/10.3758/s13423-017-1230-y, PubMed: 28251595
Shinkuma, R., Nishida, S., Kado, M., Maeda, N., & Nishimoto, S. (2019). Relational network of people constructed on the basis of similarity of brain activities. *IEEE Access*, *7*, 110258–110266. https://doi.org/10.1109/ACCESS.2019.2933990
Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception*, *29*, 1143–1154. https://doi.org/10.1068/p3104, PubMed: 11220207
Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, *11*, 54–71. https://doi.org/10.1037/1082-989X.11.1.54, PubMed: 16594767
Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11*, 13. https://doi.org/10.1167/11.5.13, PubMed: 22207654
Tallon-Baudry, C. (2022). The topological space of subjective experience. *Trends in Cognitive Sciences*, *26*, 1068–1069. https://doi.org/10.1016/j.tics.2022.09.002, PubMed: 36243671
Tsuchiya, N., & Saigo, H. (2021). A relational approach to consciousness: Categories of level and contents of consciousness. *Neuroscience of Consciousness*, *2021*, niab034. https://doi.org/10.1093/nc/niab034, PubMed: 34659799
Tsuchiya, N., Taguchi, S., & Saigo, H. (2016). Using category theory to assess the relationship between consciousness and integrated information theory. *Neuroscience Research*, *107*, 1–7. https://doi.org/10.1016/j.neures.2015.12.007, PubMed: 26748074
van Boxtel, J. J. A., Tsuchiya, N., & Koch, C. (2010). Opposing effects of attention and consciousness on afterimages. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 8883–8888. https://doi.org/10.1073/pnas.0913292107, PubMed: 20424112
Vanston, J. E., & Crognale, M. A. (2018). Effects of eccentricity on color contrast. *Journal of the Optical Society of America A*, *35*, B122–B129. https://doi.org/10.1364/JOSAA.35.00B122, PubMed: 29603965
Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the savage–dickey method. *Cognitive Psychology*, *60*, 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001, PubMed: 20064637
Wagner, G., & Boynton, R. M. (1972). Comparison of four methods of heterochromatic photometry. *Journal of the Optical Society of America*, *62*, 1508–1515. https://doi.org/10.1364/JOSA.62.001508, PubMed: 4643012
Watson, A. B. (2014). A formula for human retinal ganglion cell receptive field density as a function of visual field location. *Journal of Vision*, *14*, 15. https://doi.org/10.1167/14.7.15, PubMed: 24982468
Williams, D. R., MacLeod, D. I. A., & Hayhoe, M. M. (1981). Foveal tritanopia. *Vision Research*, *21*, 1341–1356. https://doi.org/10.1016/0042-6989(81)90241-8, PubMed: 6976039
Wolfe, J. M. (1999). Inattentional amnesia. In *Fleeting memories: Cognition of brief visual stimuli* (pp. 71–94). Cambridge, MA: MIT Press.
Yang, J., Qi, X., & Makous, W. (1995). Zero frequency masking and a model of contrast sensitivity. *Vision Research*, *35*, 1965–1978. https://doi.org/10.1016/0042-6989(94)00285-T, PubMed: 7660602
Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75. https://doi.org/10.1038/23936, PubMed: 9817201
Yeshurun, Y., & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience*, *3*, 622–627. https://doi.org/10.1038/75804, PubMed: 10816320