https://doi.org/10.2352/CIC.2024.32.1.9

Corresponding Colors in Virtual Reality: A Proof of Concept

Nicoletta Prencipe, Mikki Hiltunen, Lauri Klemettilä, Evan G. Center, Matti Pouke, Anna Yershova, Timo Ojala, Steven M. LaValle; Center for Ubiquitous Computing, Faculty of Information Technology and Electrical Engineering, University of Oulu; Oulu, Finland

Abstract

Corresponding colors data are necessary to evaluate and improve chromatic adaptation models. Virtual reality (VR) technologies offer a promising and flexible tool for studying chromatic adaptation. This work revisits the classic technique of haploscopic matching to collect corresponding colors data using VR. A virtual environment was designed in which the same visual scene was lit with a light of a different chromaticity for each eye, and an exploratory study was conducted, collecting corresponding colors data from 10 participants (5 color stimuli, 3 trials for each). The averaged CIEDE2000 standard deviation for individual observers, across 5 colors, was 2.94*, showing adequate precision despite the low amount of repetitions. This work represents a proof of concept for an efficient and realistic VR haploscopic matching paradigm, which may be extended by future studies employing display color characterization and greater numbers of participants and trials.*

Keywords: *Virtual reality - Chromatic adaptation - Corresponding colors data - Haploscopic matching.*

Introduction

The human visual system is robust with respect to changes in the illumination conditions that are close to daylight. This capability of discarding colored lights is known as chromatic adaptation, and the consequent invariance of the perceived objects' colors is usually referred to as color constancy. Many applications in both camera and display science rely on chromatic adaptation models, referring to the domain of computational color constancy [18, 5, 17]. Corresponding colors are defined as two color stimuli perceived as matching under two different viewing conditions [14]. In other words, when viewing one color while fully adapted to a specific illuminant, it appears the same as another color when fully adapted to a different illuminant. Corresponding colors data represent the benchmark to evaluate the performance of the so-called chromatic adaptation transforms (CATs) as models emulating chromatic adaptation [1, 28]. A CAT should map one corresponding color into the other. CATs are parametrized by the information concerning the two viewing conditions in one intended direction, meaning that CATs are meant to map color stimuli from the test viewing condition to the reference one. Recent studies have argued that CATs might not be reversible [10].

The search for an effective CAT is an active research line [4, 16]; due to the incomplete understanding of chromatic adaptation (both low-level sensory and high-level neural mechanisms are involved [14, 15]) and the challenges associated with collecting corresponding color data, as well as their scarcity. The main difficulty in collecting corresponding colors data lies in the fact that it is non-trivial to find a way to attain the condition of matching colors that are perceived under two different viewing conditions. Two of the main techniques used to tackle this issue are memory matching and haploscopic matching [12, 14]. Memory matching consists of letting the observer match a reference color stimulus that he/she is currently perceiving while adapted to a certain illuminant, with a test stimulus previously perceived, while adapted to a different illuminant [27, 10]. In haploscopic matching, sometimes referred to as simultaneous or binocular matching, each eye immersed in an environment lit by a different illuminant, and both test and reference stimuli are presented simultaneously [23, 8, 7, 30]. This condition is achieved by using ad-hoc physical devices, sometimes called haploscopes, which create binocular rivalry. Using memory matching requires less specific equipment, but implies some memory-induced distortions and a consequent intrinsic lack of precision, resulting in the need of increasing the number of trials for each color stimulus. On the other hand, the haploscopic technique has no memory-induced distortions, but it is based on the simplifying assumption that the two eyes adapt independently, and it does not take into account e.g. higher level brain mechanisms involved in the adaptation process. Overall, haploscopic matching is faster than memory matching because it does not require waiting for adaptation time during each trial. However, it has been used less frequently than memory matching in recent studies.

This exploratory work proposes the use of virtual reality (VR) to perform haploscopic matching experiments. The haploscopic matching technique seems particularly suitable to be replicated in VR since there is full control over what each eye sees and the observer is totally immersed in a controlled environment. As it will be detailed in the next section, VR headsets would circumvent the need for the participant to keep his/her head still in front of the two visual fields (test and reference field in classic haploscopic matching), facilitating a more natural viewing environment. Furthermore, being able to collect corresponding colors data in VR would eliminate the need for specific physical equipment as everything is simulated virtually. The availability of economic consumer-grade VR headsets would facilitate easier data collection. One clear limitation of using VR technologies for colorimetric experiments is that RGB-controlled illumination conditions are used, even though they do not correspond to the spectral power distribution of standard CIE illuminants. However the feasibility of color vision research in VR has been analyzed in recent works [19, 3, 13]. VR has been utilized in various studies concerning lighting perception [2], and color constancy in complex scenes [21, 20]. The scope of this work is to revisit and design the haploscopic matching technique in a VR context, promoting its use over the memory matching approach to be able to collect more and more precise corresponding colors data. We must stress that the current study did not involve a display color characterization [6], and calibration, which will be part of a future work.

The present work is organized as follows: first, the design of the virtual environment for haploscopic matching and the adopted experimental procedure are described; then, the VR corresponding colors data from the 10 subjects who took part to the study are presented, as well as questionnaire data collected after the color matching task; finally, future perspectives are discussed.

Virtual Environment Design

The Virtual Environment (VE) for haploscopic matching was developed with the Godot 4.2 engine using GDScript on a Windows 11 PC. The Godot engine has out-of-the-box XR support. Godot was chosen for its relative simplicity, ease-of-use, and streamlined rendering pipeline. Although more advanced rendering techniques are provided by the more popular engines, such as Unity and Unreal Engine, these techniques were not considered advantageous in this pilot study, as we wanted to minimize the number of variables affecting rendered pixel colors. The VE consisted of an empty space, with no particular depth cues. Both head translations and rotations were enabled, but locomotion in the VE was not possible, thus participants were sitting during the experiment (Figure 1). Two distinct and separated square-shaped color patches were present in the VE. The colors of the patches were implemented as Godot's spatial shader materials, so that the albedo color of the material could be defined separately for each eye. Other properties, including diffuse and specular shading models, were left undefined resulting in diffuse nonmetal materials utilizing Godot 4.2 default PBR shading models [1]. We used the linear tone mapper option with exposure and white point set to default to eliminate post-processing effects on the rendered colors. The patch on the right was always the test color stimulus, while the left one was the reference, adjustable color stimulus. They were placed at the observer's gaze height, in front of him/her and perpendicular to the ground. Nevertheless, the color patches were not anchored to the participants' gaze, meaning that they could look around and move their head without the patches following their gaze. In in-game units roughly corresponding to real-life meters, the patches were 0.64 units in height, separated by a distance of 1.36 units and placed at eye-level 10 units away from the participants point of view, resulting in patches that subtended approximately 3.7° of vertical and horizontal visual angle.

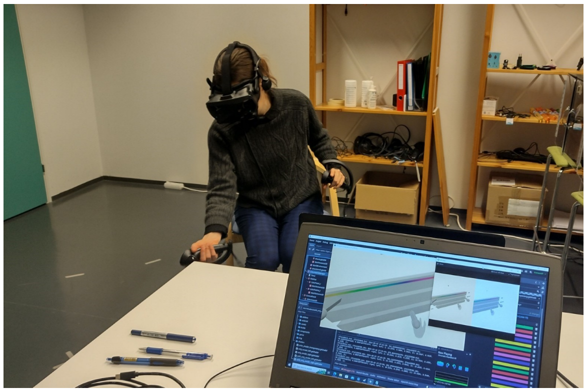

Figure 1. *A picture of the experimental setting. The participant was sitting, wearing a Valve Index headset, and interacting with the color sliders using controllers.*

The key element in the design of the VE is that each eye sees the environment as being lit by an omnidirectional light of a different color (Figure 2). In particular, the right eye always sees the environment as being lit by white light, with chromaticity of the D65 standard illuminant, while the left eye sees the environment as being enlightened by a slightly orangish light, having an intermediate chromaticity between the standard illuminants A and D50, with CIE 1931 xy coordinates $(0.3870, 0.3868)$; see Figure 4 for a plot of the illuminants' chromaticities. The lights were placed directly behind the observer and they did not cast any shadows. Global illumination was not simulated, so that the colors of the patches are not affected by the light bouncing off surrounding surfaces.

Another important factor is how this unusual lighting condition relates to the two color patches. In fact, the right eye only sees the right patch (test color stimulus), lit by the white light and sees nothing (the color of the background) where the left patch is placed, while the left eye only sees the left patch (reference color stimulus), lit by the orangish light and sees nothing where the right patch is supposed to be (Figure 2). We must stress that the two patches are neither adjacent nor overlapping, hence, when both eyes are opened, the two separate squares are visible in the VE.

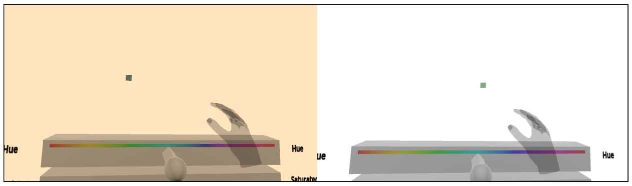

Figure 2. *Left and right displays: the left eye sees the VE as being lit by orangish light, only the left (reference) patch, lit by orangish light, is visible to the left eye; the right eye sees the same VE, but is lit by white light, and only the right (test) patch, lit by white light, is visible to the right eye. Note that the patches shown in the screenshots appear smaller than they do to the user, when wearing the headset.*

Although both the difference in illumination between the eyes and the visibility of only one of the two patches to each eye induce binocular rivalry, these features were necessary to extend the haploscopic matching technique to VR. In fact, the former allows, like in classic haploscopic matching, a visual field that differs in the illumination condition for each eye, but, in contrast to classic haploscopic matching, does not require the observer to keep his/her head still between the two viewing fields for the entire duration of the experiment [7]. With this design the discrepancy in the illumination is somehow observer-fixed, and the observer is free to move his/her head. Because of this, the latter (namely the visibility of only one patch to each eye) ensures that, as in classic haploscopic matching, only the reference (left) stimulus is present in the reference visual field (hence the one shown to the left eye), and only the test (right) stimulus is visible in the test visual field (shown to the right eye). Both aspects rely on the simplifying assumption in haploscopic matching that the eyes adapt independently and do not take into account neural mechanisms occurring in chromatic adaptation. The question about how binocular rivalry affected the user's experience was treated in the exploratory study via a questionnaire (see the section 'Questionnaire data'). We were also concerned that eye dominance might

---

[1] https://docs.godotengine.org/en/stable/classes/class_material.html

play a privileged role when the two eyes are simultaneously exposed to two different illumination conditions, thus this aspect will be discussed as well in the following sections.

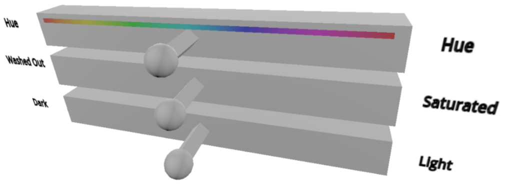

Figure 3. *The sliders that the participants could manipulate using controllers in order to adjust the color of the left (reference) patch. Sliders were chosen to be achromatic, but, to help participants that were not familiar with color spaces, intuitive indications were added, such as as a colored strip with the hue gradient for the hue and some text at the extremities of the sliders: 'Hue' for the hue, 'Washed Out' and 'Saturated' for the saturation, and 'Dark' and 'Light' for the value.*

As in classic haploscopic matching, the test color stimulus (patch on the right, visible only to the right eye) stays fixed, while the reference color stimulus (patch on the left, visible only to the left eye) is adjusted by the observer until the two stimuli match in appearance. The need for color atlases, used in classic haploscopic and memory matching experiments to pick the reference color stimulus, was circumvented by letting the observer perform adjustments using virtual sliders. The color of the right patch (before being lit by the orangish light) was controlled by the participant via three horizontal sliders (Figure 3). The sliders were positioned close to the observer in the same direction of the color patches, but at a lower height in order to avoid occlusion and to be easily accessible while sitting (Figure 1). Observers could manipulate the sliders using both controllers, while see-through hands were displayed in the VE (Figure 2). The sliders' position could be adjusted for each participant before the experiment began. HSV color coordinates were chosen for the sliders as intuitive and simple. In the implementation Godot's built-in color conversion functions from HSV to sRGB were used. Furthermore, it has been shown that the HSV color solid seems to have better navigability properties than the more perceptually optimized IPT color space [11].

Experimental Procedure

The title of the study to which the participants enrolled was 'Color matching in VR'. Subjects who reported to have color vision deficiencies were not permitted to enroll. Informed consent was obtained from all participants prior to participation in the study and all procedures were approved by the University of Oulu's ethical review board. Before the participant read and signed the consent forms, the experimenter explained that the experiment was about color perception in VR, and that the participant would wear the headset and interact with the environment using controllers while sitting. They would see two color patches and change the color of one of them using three sliders until they both looked the same. Then, the meaning of each slider was explained using the picture in Figure 3. The participant had to perform 15 matching tasks with no time limitation. Whenever the subject was satisfied with his/her match, he/she had to say 'Next' to the experimenter who would let him/her move to the next matching task.

After signing the consent forms, participants took the Ishihara test for color blindness (12 plates) to confirm the absence of color vision deficiencies. They were then tested for eye dominance using the Miles technique [29]. We use the term 'dominant manipulating' to refer to participants who were found to be left-eye dominant, meaning that the reference stimulus (the adjustable color patch) was visible to their (dominant) left eye and lit by the orangish illuminant for the entire duration of the experiment, and 'non-dominant manipulating' to refer to participants who were found to be right-eye dominant, meaning that the reference stimulus (the adjustable color patch) was lit by orangish light and displayed to their (non-dominant) left eye. Subsequently, they completed a baseline simulator sickness questionnaire (SSQ; [25]).

A Valve Index head-mounted display (HMD) was chosen for the experiment. It has two LCD screens with a resolution of 1440 × 1600 pixels per eye, a refresh rate of 120 Hz, 108° horizontal and 104° vertical field of view, and dual-element canted Fresnel lenses. The device is glasses compatible, hence participants having corrected vision were not excluded. The wide field of view of this HMD was suitable to facilitate immersion and adaptation. The software was displayed in the headset via SteamVR.

After ensuring that the participant was sitting in the right position, and wearing headset and controllers correctly, the experimenter started a timer for three minutes. During this time, everything but the two patches was present in the VE. While the observer adapted to the illumination described in the previous section, the experimenter asked him/her to become familiar with moving the virtual sliders. While the usual time needed to reach full adaptation is generally two minutes [14], we added one extra minute because of the unusual viewing condition. After the timer was up, the first matching task started.

We chose five test colors with chromaticities resembling to the ones of the 13-17th colors in the Macbeth ColorChecker (Table 3, Figure 4); in the following we will refer to them as blue (B), green (G), red (R), yellow (Y), and magenta (M). The matching of each test color was repeated 3 times, for a total of 15 matching tasks. In the prototyping phase, the choice to have only 3 repetitions for each test color was made to ensure the participants' comfort. We found that 15 matching tasks were a tolerable workload, especially for individuals not accustomed to wearing a VR headset. As detailed later, this did not significantly affect intra-observer variability. The 15 test colors appeared in a shuffled order (the same for all subjects: RMBY|RBGMY|BRYMG), ensuring that no test color appeared consecutively.

The duration of each matching task, as well as the overall time spent in VR (including adaptation time), was recorded, along with the selected matching reference color stimuli. The picked reference color stimuli (being lit by the orangish light) as well as the test color stimuli and the illuminants' colors, were saved as sRGB outputs of the Godot's rendering engine, obtained via screenshots of Steam VR, and then converted to XYZ (Table 3). This means that the preliminary results presented in the following section were obtained from the sRGB outputs of the rendering engine, without taking into account the HMD's displays color characterization and without prior calibration of the HMD (see the section 'Conclusion and future work' for future improvements on this aspect). After completing the matching tasks, the experimenter administered a second questionnaire composed of two parts: questions about the participant's overall experience, and again the SSQ to test for any potential symptoms of cybersickness induced by the task. Finally, the experimenter answered any questions and rewarded the subjects with University of Oulu merchandise, worth approximately 5 euros. The overall duration of the experiment was about 30 minutes.

Preliminary Results

The study involved 10 participants (two of them identified as female), with a mean age of 29.1 years and range of 21-42 years. When asked about previous experience playing computer games, two reported playing every day, one several times per week, two once or twice per week, one once or twice per month, two once or twice per year, one just once or twice ever, and one had never played. Regarding previous VR experience, one reported using VR once or twice per week, one once or twice per month, four once or twice per year, and four just once or twice ever (thus all had tried VR at least once before).

Concerning eye dominance, 4 participants were left-eye dominant (dominant manipulating), and 6 right-eye dominant (non-dominant manipulating). As determined by graphical inspection and Shapiro-Wilk tests, distributions for standard deviations, durations, and SSQ total scores were all non-normal and thus non-parametric statistical tests were used. The average time spent by participants in VR (including adaptation time and 15 matching tasks) was 17.3 min, with a minimum of 11 min (observer 1) and a maximum of 22.5 min (observer 8). Table 1 shows the average duration of the matching task for each test color stimulus. A Friedman test showed no significant difference in matching durations among colors, $\chi^2(4, N = 10) = 5.12, p = .28$, $W = .13$. The mean time of the dominant manipulating subjects was 0.86 min, while the mean time of the non-dominant manipulating subjects was 0.95 min.

Table 1. Average duration of the matching task for each test color stimulus, across all trials and participants.

	Blue	Green	Red	Yellow	Magenta
mean duration (min)	0.87	0.92	0.78	1.10	0.89

CIE 1931 XYZ coordinates of the illuminants, test color stimuli and selected reference color stimuli (averaged in CIE L*a*b* across all the trials and all participants) are reported in Table 3 and plotted in the CIE 1976 $u'v'$ chromaticity diagram in Figure 4. Note that, as in classic corresponding colors data, the plot in Figure 4 depicts an uneven shift from the matching stimuli to the test stimuli, in the direction of the orangish illuminant.

In order to evaluate intra-observer variability, the standard deviation (SD) across the 3 trials for each observer and each test color was computed using CIEDE2000 color metric, as in [9, 10]. The SD values are reported in Table 4 and plotted in Figure 5. A Friedman test showed a significant difference in standard deviations among colors, $\chi^2(4, N = 10) = 11.80$, $p = .019$, $W = .29$, with yellow being particularly precise relative to the other colors. Low levels of precision for the red test color stimulus might have been affected by the fact that the red hue was placed at the extremities of the hue slider (see Figure 3). SDs for individual observers range from 0.80 to 8.97, with a mean of 2.94 CIEDE2000 units, which provides an acceptable level of precision despite having only three repetitions per color.

Table 3: CIE XYZ coordinates of illuminants, test color stimuli and selected reference color stimuli. The latter were obtained by averaging the selected reference color stimuli in CIE L*a*b* coordinates across all trials and participants.

	Test color			Matching reference color		
Color	X	Y	Z	X	Y	Z
Illuminant	0.9505	1.0000	1.0890	0.7078	0.7075	0.4138
Blue	0.1204	0.1099	0.2478	0.1085	0.1039	0.1381
Green	0.1491	0.2012	0.1345	0.1567	0.2004	0.0766
Red	0.1835	0.1431	0.1124	0.1854	0.1381	0.0584
Yellow	0.3741	0.4017	0.1356	0.3403	0.3407	0.0785
Magenta	0.2370	0.1847	0.2560	0.2285	0.1836	0.1444

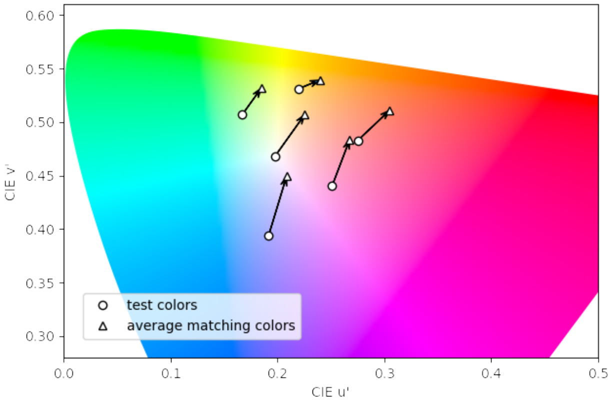

Figure 4. *Color stimuli of Table 3 plotted in CIE 1976 $u'v'$ chromaticity diagram.*

Table 4: Standard deviation (SD), calculated using CIEDE2000 color difference, for the 3 trials of each of the 10 subjects.

Subject	Blue	Green	Red	Yellow	Magenta	**Mean**
1	3.32	2.61	3.63	1.85	5.94	**3.47**
2	2.40	2.13	1.65	1.82	1.70	**1.94**
3	5.20	2.77	2.96	1.97	4.44	**3.47**
4	2.02	0.80	2.15	1.41	1.99	**1.68**
5	8.97	3.10	1.87	1.89	4.69	**4.10**
6	4.27	4.15	3.63	1.34	2.60	**3.20**
7	2.28	2.58	4.89	2.08	3.77	**3.12**
8	4.07	1.66	2.24	1.80	1.75	**2.30**
9	1.75	3.32	3.31	2.51	4.46	**3.07**
10	2.80	2.49	4.15	1.37	4.37	**3.04**
Mean	**3.71**	**2.56**	**3.05**	**1.80**	**3.57**	**2.94**

In order to evaluate inter-observer variability, the SD using CIEDE2000 of overall observers was calculated for each color. The values are reported in Table 5. Typical inter-observer variability in chromatic adaptation is around 4 CIEDE2000 units, [1, 10]. One can see that the average SD of the overall data is 4.37, which is consistent with 4 CIEDE2000 units.

Let us now return to the issue of eye dominance in our setup,

Table 5. SD of overall participants for each test color stimulus, calculated using CIEDE2000.

	Blue	Green	Red	Yellow	Magenta	**Mean**
overall SD	6.20	3.93	4.21	2.56	4.96	**4.37**

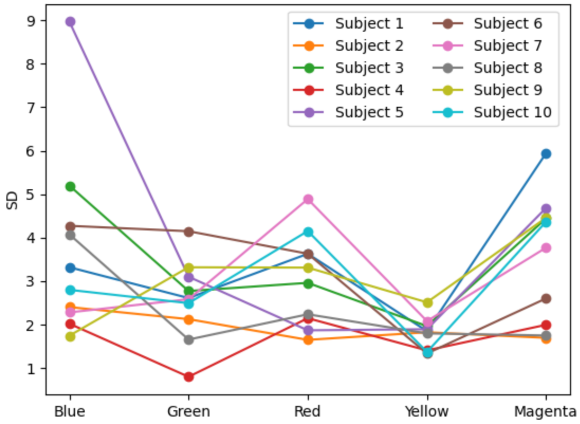

Figure 5. *Plot of the SD across the 3 trials, for each subject and each test color stimulus.*

where the reference and test patches were yoked to the left and right eyes, respectively. Standard deviations for right-eye dominant participants ($Mdn = 3.33$) were significantly greater than those for left-eye dominant participants ($Mdn = 2.12$) as determined by a Wilcoxon ranked-sum test ($Z = 2.07$, $p = .038$, $r = 0.66$). This difference in precision was present despite no significant difference in the average judgement duration taken between right-eye dominant ($Mdn = 0.95$ min) and left-eye dominant ($Mdn = 0.82$ min) participants ($Z = 0.30$, $p = .76$, $r = 0.10$). A possible explanation could be that the eye which sees the adjustable patch has a higher workload in this experiment. In fact, the reference stimulus varies more often than the test stimulus (fixed during each matching task) as it is subject to many subsequent small adjustments, thus it seems reasonable that the preferred eye could deal better with a high workload than the non-dominant eye. This means that precision of corresponding colors data could be improved by performing haploscopic matching experiments so that the reference stimulus is displayed to the subject's dominant eye.

Questionnaire Data

When asked to rate the general difficulty of the color matching task on a scale from 1 (very easy) to 7 (very difficult), our sample tended to find the task neither too easy nor too difficult ($Mdn = 4$; $range = 2 - 6$). Consistent with precision data (SDs; Table 4 and Figure 5), yellow was the color most often subjectively identified as the most challenging to match (selected 3 times; red: 0, green: 2, blue: 1, magenta: 2, all equal: 4). Likewise, yellow was also the color least often subjectively identified as easiest to match (selected 1 time; red: 3, green: 2, blue: 2, magenta: 3, all equal: 4). As reported in Table 1, the average duration for matching the yellow test stimulus was the maximal one. One possible explanation for the difficulty in matching the yellow test stimulus could be its proximity to the color of the orangish illuminant.

We suspected that the binocular rivalry created by forcing participants to fuse different images presented to each eye might result in some form of cybersickness (e.g., [22]), particularly in regard to symptoms related to ocular strain. Therefore, we administered the simulator sickness questionnaire (SSQ [25]) prior to and after the VR exposure to address this possibility. Due to a procedural oversight, the first three participants did not complete the pre-VR SSQ, hence only data from the remaining seven participants with both pre- and post-VR SSQ data are statistically compared here. SSQ total scores after VR exposure ($Mdn = 14.96$; for $n = 10$, $Mdn = 24.31$) were greater than those before VR exposure ($Mdn = 3.74$). However, this difference did not surpass statistical significance as indicated by a Wilcoxon-signed rank test ($Z = 1.89$, $p = .058$, $r = 0.72$). This trend appeared to be more driven by the ocular ($Z = 1.79$, $p = .074$, $r = 0.68$) and nausea ($Z = 1.66$, $p = .098$, $r = 0.63$) subcomponents of the SSQ than the disorientation ($Z = 1.10$, $p = .27$, $r = 0.42$) subcomponent. We cannot draw too strong of conclusions from this small sample, though if similar effect sizes are revealed in future work, we must consider the implications that cybersickness may carry for simultaneous color matching paradigms.

We were curious whether participants would notice that different images were shown to each eye. At the end of the questionnaire, participants were asked directly whether they tried closing one eye and looking only with the other. Five participants reported having tried viewing with only one eye and five reported using both eyes throughout. Individual SDs did not statistically differ between participants who reported trying to use one eye at least once ($Mdn = 3.12$) and those who did not ($Mdn = 3.07$), ($Z = 0.25$, $p = .42$, $r = 0.25$). When participants were asked directly whether they noticed 'anything strange' about the VE, two said no, five mentioned details about one patch displaying to only one eye, and three mentioned details regarding general blurring, fading, or other artifacts. Interestingly, none explicitly mentioned any differences in illumination between the two eyes. Anecdotally, participants reported, consistent with [24, 11], that they used the hue slider more often than the other sliders, which were used only for fine-tuning.

Conclusion and Future Work

Recent advancements in VR technologies present a promising opportunity to introduce innovative solutions to the issues related to collecting corresponding colors data. In this work, we proposed a novel haploscopic matching experiment in VR. By design this experimental technique does not require the observer to keep his/her head in a fixed position for the whole duration of the experiment, as in classic haploscopic matching. The preliminary results obtained in this study are comparable with their state-of-the-art analogues, especially in terms of the individual SD, despite the low amount of trials collected for each color. Nevertheless, we must stress that, in order to establish this novel VR haploscopic matching technique as a new paradigm to collect more reliable and precise corresponding colors data, it is of fundamental importance in future studies to involve the HMD's displays color characterization [6] and calibration, as suggested by recent research on color fidelity and colorimetric accuracy in VE [19, 13, 3]. It would be of great interest to compare datasets collected in this way with the predictions of different CATs. Furthermore, for future studies, different shuffled order of test colors for each participant (while maintaining no consecutive colors) should be adopted (to average out possible afterimage or order effects), and the number of repetitions needs to be increased for better comparison with other corresponding colors datasets and to potentially improve precision. However, the number of trials is somewhat constrained by the limited time a person can effectively spend in VR matching colors. Increasing the number of trials significantly might necessitate reducing the number of colors or dividing the experiment into multiple sessions. Insights from the average duration of a matching task (Table 1) could assist in making informed decisions, when designing future studies. Furthermore, the hue gradient could be removed from the hue slider to prevent any potential influence on the matching task. Alternative color spaces could be explored for the sliders, as in [26], along with the possibility of zooming in on a region of the slider to obtain more precise measurements. Eye dominance seems to play a role that resulted to be statistically relevant already with 10 participants, implying that observers seeing the reference adjustable stimulus with their dominant eye obtained more precise and consistent matches. Finally, increasing the number of subjects could benefit future experiments in numerous ways, such as identifying potential correlations between duration and test color stimuli (e.g., the case of the yellow stimulus) and potentially enabling conclusions about induced cybersickness and its relation to color matching performance, if any such relationship exists. Despite carrying some minor limitations, we identify haploscopic color matching in VR as an efficient, and potentially, robust, method of collecting new corresponding colors datasets.

Acknowledgments

This research was conducted with the financial support of Infotech Oulu and the European Research Council Advanced Grant (ERC AdG, ILLUSIVE: Foundations of Perception Engineering, 101020977).

References

[1] CIE TC 1-52. CIE Publ. 160 A review of chromatic adaptation transforms. 2004;160:29.

[2] A. Bellazzi, L. Bellia, G. Chinazzo, F. Corbisiero, P. D'Agostino, A. Devitofrancesco, F. Fragliasso, M. Ghellere, V. Megale, and F. Salamone. Virtual reality for assessing visual quality and lighting perception: A systematic review. *Building and Environment*, 209:108674, 2022.

[3] U. Bhaumik, F.B. Leloup, and K. Smet. Systematic comparison of head mounted display colorimetric performance using various color characterization models. *Optics Continuum*, 2(6):1490–1504, 2023.

[4] S. Bianco and R. Schettini. Two new von Kries based chromatic adaptation transforms found by numerical optimization. *Color Research & Application*, 35(3):184–192, 2010.

[5] S. Bianco and R. Schettini. Computational color constancy. In *3rd European Workshop on Visual Information Processing*, pages 1–7. IEEE, 2011.

[6] D.H. Brainard, D.G. Pelli, and T. Robson. Display characterization. *Signal Process*, 80:2–067, 2002.

[7] E.J. Breneman. Corresponding chromaticities for different states of adaptation to complex visual fields. *JOSA A*, 4(6):1115–1129, 1987.

[8] R.W. Burnham, R.M. Evans, and S.M. Newhall. Prediction of color appearance with different adaptation illuminations. *JOSA*, 47(1):35–42, 1957.

[9] S. Cai and M.D. Fairchild. Individual corresponding colors data and chromatic adaptation transforms. In *Color and Imaging Conference*, volume 24, pages 248–253. Society for Imaging Science and Technology, 2016.

[10] S. Cai and M.D. Fairchild. Bidirectional individual corresponding colors data. *Color Research & Application*, 43(5):643–654, 2018.

[11] T.D. Canham, I.S. MacKenzie, R.F. Murray, and M.S. Brown. The effect of perceptual optimization on color space navigability. In *Graphics Interface 2023*, 2022.

[12] M.D. De Fez, P. Capilla, M.J. Luque, J. Pérez-Carpinell, and J.C. Del Pozo. Asymmetric colour matching: Memory matching versus simultaneous matching. *Color Research & Application*, 26(6):458–468, 2001.

[13] F. Díaz-Barrancas, R. Gil Rodríguez, F.S. Bayer, A. Aizenman, and K.R. Gegenfurtner. High-fidelity color characterization in virtual reality across head mounted displays, game engines, and materials. *Optics Express*, 32(13):22388–22409, 2024.

[14] M.D. Fairchild. *Color appearance models*. Wiley, 2013.

[15] M.D. Fairchild. What sets the state of chromatic adaptation? *Journal of Vision*, 19(15):40–40, 2019.

[16] M.D. Fairchild. Von Kries 2020: Evolution of degree of chromatic adaptation. In *Color and Imaging Conference*, volume 28, pages 252–257. Society for Imaging Science and Technology, 2020.

[17] G.D. Finlayson and S.D. Hordley. Color constancy at a pixel. *JOSA A*, 18(2):253–264, 2001.

[18] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational color constancy: Survey and experiments. *IEEE transactions on image processing*, 20(9):2475–2489, 2011.

[19] R. Gil Rodríguez, F. Bayer, M. Toscani, D. Guarnera, G.C. Guarnera, and K.R. Gegenfurtner. Colour calibration of a head mounted display for colour vision research using virtual reality. *SN Computer Science*, 3:1–10, 2022.

[20] R. Gil Rodríguez, L. Hedjar, M. Toscani, D. Guarnera, G.C. Guarnera, and K.R. Gegenfurtner. Color constancy mechanisms in virtual reality environments. *Journal of Vision*, 24(5):6–6, 2024.

[21] R. Gil Rodríguez, M. Toscani, D. Guarnera, G.C. Guarnera, F. Bayer, and K.R. Gegenfurtner. Color constancy in a virtual reality environment. *Journal of Vision*, 20(11):1226–1226, 2020.

[22] J. Häkkinen. 61.3: Invited paper: A virtual display for mobile use. In *SID Symposium Digest of Technical Papers*, volume 35, pages 1586–1589. Wiley Online Library, 2004.

[23] R.W.G. Hunt. Light and dark adaptation and the perception of color. *JOSA*, 42(3):190–199, 1952.

[24] D.B. Judd. Ideal color space. *Color Eng*, 8(44):2405, 1970.

[25] R.S. Kennedy, N.E. Lane, K.S. Berbaum, and M.G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[26] J.M. Kraft and D.H. Brainard. Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences*, 96(1):307–312, 1999.

[27] K.M. Lam. *Metamerism and Colour Constancy*. PhD thesis, University of Bradford, Bradford, UK, 1985.

[28] M.R. Luo. A review of chromatic adaptation transforms. *Review of Progress in Coloration and Related Topics*, 30:77–92, 2000.

[29] W.R. Miles. Ocular dominance in human adults. *The journal of general psychology*, 3(3):412–430, 1930.

[30] L. Mori, H. Sobagaki, H. Komatsubara, and K. Ikeda. Field trials on CIE chromatic adaptation formula. *CIE Proceeding, 22nd Session*, pages 55–58, 1991.