This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3195111

Analysis of the Saliency of Color-Based Dichoptic Cues in Optical See-Through Augmented Reality

Austin Erickson, *Graduate Student Member, IEEE,* Gerd Bruder, *Member, IEEE,* and Gregory F. Welch, *Fellow, IEEE*

Abstract—In a future of pervasive augmented reality (AR), AR systems will need to be able to efficiently draw or guide the attention of the user to visual points of interest in their physical-virtual environment. Since AR imagery is overlaid on top of the user's view of their physical environment, these attention guidance techniques must not only compete with other virtual imagery, but also with distracting or attention-grabbing features in the user's physical environment. Because of the wide range of physical-virtual environments that pervasive AR users will find themselves in, it is difficult to design visual cues that "pop out" to the user without performing a visual analysis of the user's environment, and changing the appearance of the cue to stand out from its surroundings.
In this paper, we present an initial investigation into the potential uses of dichoptic visual cues for optical see-through AR displays, specifically cues that involve having a difference in hue, saturation, or value between the user's eyes. These types of cues have been shown to be preattentively processed by the user when presented on other stereoscopic displays, and may also be an effective method of drawing user attention on optical see-through AR displays. We present two user studies: one that evaluates the saliency of dichoptic visual cues on optical see-through displays, and one that evaluates their subjective qualities. Our results suggest that hue-based dichoptic cues or "Forbidden Colors" may be particularly effective for these purposes, achieving significantly lower error rates in a pop out task compared to value-based and saturation-based cues.

Index Terms—Augmented Reality, Optical See-Through Displays, Visual Perception, Attention Cues, Preattentive Cues, Human-Computer Interaction (HCI)

1 Introduction

As augmented reality technology (AR) advances, we move towards a future of pervasive AR [1], where people will wear AR displays for longer continuous periods of time and will regularly interact with AR imagery in a manner similar to how we currently interact with apps on our smart phones. Pervasive AR users will be accompanied by their personal AR imagery that is relevant to their current interests and context, as well as shared imagery from colleagues and friends, and public imagery in the form of art and advertisements. Some of this virtual imagery will be registered to physical positions in the user's environment while others may be anchored to points on the user's body or presented in the form of a heads-up display.

As the user goes about their day, their AR system will intermittently employ various visual, audio, haptic, or multimodal notifications in order to draw the user's attention to new information. This information could be something as casual as letting the user know that they have a new message to something more urgent requiring the user's immediate attention, such as a public safety alert. The cues chosen to convey the notification should be effective at quickly drawing the user's attention to the new information, however the choice of visual cues to use is not as straightforward as it may be on other types of displays.

AR displays are unique in that the imagery being displayed by the device is superimposed over the user's view of their physical environment, so any visual cues that the AR system employs to guide or draw the user's attention must contend not only with the other virtual imagery being displayed, but also with the appearance of the user's physical environment. Further, it is probable that if a generic traditional visual cue is chosen, then there exists physical-virtual environments in which it will blend in and be more difficult for the user to notice.

One potential solution for this is to make use of dichoptic visual cues for drawing attention to virtual imagery within the user's field of view. These cues involve having a discrepancy in the appearance of the cue between the user's eyes, such as by changing the color, size, or positioning of the cue in one eye. These cues are not commonly observed in daily life, with the exception of the parallax effect that we use to judge distances, and so they may be particularly effective at standing out to the user under a wide range of environment conditions. Additionally, these types of cues have an advantage in that they can draw the user's attention to virtual imagery without having to completely change the appearance of the imagery for both of the user's eyes. Instead, one eye may see the original intended image while the other eye sees a version that has been selectively altered in order to draw the attention of the user.

In this paper, we present an initial investigation into the potential uses of dichoptic visual cues for optical see-through (OST) AR displays, specifically cues that have a difference in hue, saturation, or value between the user's eyes. Some of these cues have been investigated previously, where

- *University of Central Florida, Orlando, FL, 32816.*
  E-mail: ericksona@knights.ucf.edu, gerd.bruder@ucf.edu, welch@ucf.edu

it has been shown that saturation-based, or luster cues, in addition to a subset of value-based dichoptic cues, DeadEye monocular cues, are effective as preattentive cues [2], [3]. However, their effectiveness as preattentive cues specifically on OST displays remains to be investigated, and since OST AR displays are prone to reduced contrast [4] compared to other displays and effects such as color blending [5], the effectiveness of these cues may be somewhat reduced compared to other more-traditional displays.

We present two studies: one that evaluates whether dichoptic visual cues consisting of either hue, saturation, or value discrepancies between the user's eyes are effective as preattentive cues on a particular OST display, the Microsoft HoloLens 2, and one that evaluates the subjective qualities of these types of cues. The effectiveness of the dichoptic cues is first evaluated in a pop-out style task in experiment one, where our results indicate that hue-based dichoptic visual cues significantly outperform both saturation-based and value-based dichoptic cues on the HoloLens 2. Three subjective qualities of the dichoptic cues are investigated in experiment two: the level of noticeability, the sense of implied urgency, and the level of visual comfort. These qualities are measured via Likert scale responses, and our results indicate that as the intensity level of the dichoptic cue increases, the sense of implied urgency significantly increases and the level of comfort significantly decreases.

The remainder of this paper is structured as follows: Section 2 provides background on the existing research on saliency of visual cues and their intersection with AR and virtual reality (VR) research. Section 3 describes an experiment designed to compare the saliency of color-based dichoptic cues on an OST AR display. Section 4 describes the results of experiment one. Section 5 describes a second experiment that evaluates the subjective qualities of color-based dichoptic preattentive cues, and section 6 describes the results. Section 7 provides a discussion on the implications of the results of the two studies, and section 8 concludes the paper.

2 BACKGROUND

In this section, we present an overview on saliency of visual cues, dichoptic cues, and their intersections with AR and VR research.

2.1 Saliency of Visual Cues

Visual cues are commonly used to either *direct* or *attract* the attention of an observer to something specific in their environment. The effectiveness of such cues is referred to as *saliency*, the ability of a visual cue to stand out from its surroundings. Saliency is affected by many different factors, which can be separated out into two main groups: features of the cue itself, and features in the environment near or surrounding the cue. Kamkar et al. provides an insightful overview of how these factors affect the saliency of visual cues [6].

Within these two groups are a multitude of different features which affect the saliency of visual cues in different manners. In general, a cue can be made more salient by changing its appearance to differ from the environment's appearance in one or multiple ways [7]. Color, form, motion, and positioning changes have been shown in the past to be particularly effective at increasing cue saliency [8]. Increasing the number features that differ between a cue and its surroundings has also been shown to further increase the saliency of the cue, for example combining a luminance-based cue with size-based cue, or a color-based cue with a motion-based cue [9], [10].

The saliency of a visual cue can be measured in several different manners, such as through performance analysis of search tasks performed by participants, or through subjective measures. For search tasks, there are two main varieties: feature search and conjunction search. In feature search, sometimes referred to as a pop-out task, the participant is tasked with identifying an object within an arrangement of distractor objects that differs from the others in one or more certain distinct features [11]. Participant performance in such tasks has been shown to be unaffected by increasing the number of distractor objects if the cue is successful at "popping-out" to the user [12]. In conjunction search, participants are similarly tasked with identifying an object with certain features, however in this task the distractor objects share a subset of the features with the object the participant is searching for [11], [13]. Participant performance in this type of task has been shown to be significantly reliant on the number of distractor objects presented, implying that participants must perform a serial search of all objects to find the one with the specified features [11].

2.2 Dichoptic Visual Cues

Dichoptic visual cues are cues that appear differently between each of the observer's eyes, creating a phenomenon known as binocular rivalry [14]. Interestingly, such cues may be perceived differently over time by the observer, where at the initial onset of the rivalry, the observer's initial perception may be somewhat stable until alternation starts to occur, where the appearance of the cue seems to alternate as the image from one eye dominates while the other is suppressed [15], [16]. This difference in image shown between eyes can be in the form of hue, lightness, size, positioning, or any combination of the four. Wolfe and Franzel studied these types of cues in a series of experiments in 1988 [2]. They specifically investigated form rivalry, color rivalry, and binocular luster, where binocular luster is achieved by having a disparity in the perceived lightness of the visual cue such that one eye sees a darker cue while the other sees a lighter cue. Of these techniques, only binocular luster was found to successfully pop-out to study participants, however they concluded that binocular rivalry may be an effective manner of guiding the attention of an observer.

Dichoptic color rivalry, sometimes referred to as "forbidden colors," were investigated by several other researchers since 1983 [17], [18], [19]. Such colors were thought to form when opposing colors, such as red and green, or blue and yellow, blended together on one or both of the observer's eyes, resulting in a color that appears as both of the input colors rather than a blending of the two [17], [18]. However in 2006, Hsieh and Tse [19] suggested that the term "forbidden" is misleading and that the combined color is actually a blending of the two colors. Whether it is a blending or not,

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3195111

such color combinations between the observer's eyes are not commonly seen, and thus may be an effective method of drawing the attention of an observer, even if they are not necessarily processed preattentively, as shown by Wolfe and Franzel [2].

More general dichoptic cues were revisited by Zou et al. in 2017, where they confirmed that such cues may be effective at guiding the attention of an observer, but they achieved results that suggested that the luster effect did not pop out as strongly as originally thought in the 1988 experiments by Wolfe and Franzel [20].

More recently, Krekhov and Krüger investigated the potential of using monoscopic variations of visual cues in order to attract the attention of the observer in a technique they called DeadEye [3]. Their technique involved removing the image of the visual cue in one of the observer's eyes to cause binocular rivalry, and when tested in a pop-out task with varying numbers of distractor objects, they demonstrated that this type of cue successfully pops-out to observers in both feature search tasks with homogeneous or heterogeneous distractors, and thus can be processed preattentively by the observer. Their initial work was performed on a stereoscopic 3D flat-panel display, but was later replicated on an immersive VR display, where similar results were achieved [21].

It is also important to consider how the effectiveness of dichoptic cues varies in response to factors besides its appearance, such as the eye dominance of the observer and the appearance of the background behind or surrounding the visual cues.

Eye dominance can be tested through several different methods, including sensory ocular dominance testing, in which a stimulus with binocular rivalry is used to detect the dominant eye, and sighted ocular dominance testing, where the observer aligns their view of a target through a hole [22]. Interestingly, these different testing methods are not always correlated [23]. Eye dominance was investigated in 2006 by Shneor et al. where they concluded that the dominant eye determined through repeated sighted ocular dominance testing has priority in visual processing tasks, which they demonstrated via increased task performance in a pop-out style search task [24]. They later showed that this increase in performance for the dominant eye extends to conjunction searches as well [25]. More recently, the impact of eye dominance on user performance in using monocular displays was investigated by Bayle et al. where they found that only four of their 18 participants exhibited better tracking performance of a monocular visual cue with one eye compared to the other [26]. However, in the investigation of DeadEye cues by Krekhov and Krüger, they did not find any significant effect of eye dominance, tested via sighted dominance testing, on user performance [3]. In a study by Browne et al. involving user performance in using a monocular display in a flight simulator environment, eye dominance tested through sighted dominance testing was again shown to not have a significant effect on user performance [27]. It appears as though there has yet to be a consensus on whether or not eye dominance is an important factor in determining user performance in applications involving monocular cues, which are similar to the conditions involved when identifying dichoptic visual cues. As such, we include eye dominance, determined through sighted dominance testing, as an independent variable in experiment one.

It is also possible that the appearance of the background behind and surrounding the dichoptic cues could influence the observer's ability to distinguish the cue. Although little work has been done to investigate this specifically for dichoptic cues, there has been several studies that have investigated similar effects that the background appearance can have on user performance when using monocular displays. This was investigated by Grudin in 2002, where he found that dynamic moving backgrounds were detrimental to user performance in using a monocular display to perform a look-up task [28]. He also showed that a visually complex static background had similar, albeit lesser, effects on user performance, and concludes that monocular displays may not be well suited for dynamic or complex backgrounds. However, this was also evaluated in 2021 by Bayle et al. where it was found that high spatial frequency physical backgrounds led to increased user performance in a tracking task involving a monocular AR display compared to conditions with a low spatial frequency background [26]. This lack of consensus on the effects of background appearance on user performance in using dichoptic visual cues led us to include it as an independent variable in experiment one.

2.3 Mixed Reality Head-Worn Displays

In the works described above, the experiments were carried out using a variety of different displays to present the visual cues and distractor objects to the observer, however very little work has been done using AR or VR displays. These types of displays present a set of unique perceptual issues in the way that imagery is presented to the user, causing effects such as vergence-accommodation conflict [29], reduced field of view [30], reduced visual acuity [31], and color issues [32], [33].

AR falls under the larger umbrella term of mixed reality [34], and according to Azuma, involves registering and superimposing interactive virtual imagery over the user's view of their real physical environment in a manner that makes it appear as though the virtual imagery is in the environment with them [35]. AR displays largely fall into two main categories: optical see-through (OST) displays and video see-through (VST) displays. With OST displays, the user sees their physical environment through a clear visor or lens and virtual imagery is superimposed through the use of devices such as beam-splitters or optical waveguides [36], [37]. On the other hand, VST displays are similar to VR displays in that they typically involve a near-eye OLED or LCD display, then the user's view of physical environment is mediated to them via external cameras [37]. In both cases, the imagery displayed on the AR display is directly affected by the appearance of the user's physical environment, and factors such as lighting conditions [38], [39] and environment color [40], [41] can impact the appearance of virtual imagery shown to the user.

OST AR displays present additional problems compared to other AR displays [4], in that the light being emitted from the display blends with the light present in the user's physical environment, causing effects such as reduced contrast in bright outdoor lighting conditions [4], and color blending [5], [32], [42], [43], where the intended color of a virtual image is shifted due to the appearance of the physical environment behind it.

Several studies have examined attention guidance in VR over the past several years, for example the specific context of cinematic VR was explored by both Rothe et al. and Nielsen et al. [44], [45]. Rothe et al. created a taxonomy of attention guidance methods for this specific context, whereas Nielsen et al. investigated the use of firefly-like objects to guide user attention in cinematic VR. Lange et al. took the firefly approach a step further by investigating the potential of using swarms of insect-like objects to guide user attention in VR, where they found that this technique outperformed several other attention guidance techniques and was perceived as having the smallest negative impact on a user's sense of immersion in the virtual experience [46]. For video see-through AR, Orlosky et al. investigated different visualization methods for combining visual and thermal information. They found that a static noise style visualization, in which temperature data was visualized based on the intensity of a static effect superimposed over the user's view of the scene, led to increased user performance in search tasks relying on temperature information compared to having no temperature visualization and compared to a dichoptic luster/saturation-based visualization technique, in which temperature information was conveyed based on the amount of binocular rivalry introduced between the user's eyes [47].

Dichoptic visualizations for stereoscopic displays were briefly introduced and investigated by Zhang et al. where seven different visualizations are implemented, several of which involve color differences between the observer's eyes, and one of which investigates monoscopic presentation of certain image features similar to the DeadEye technique [48]. They describe how these visualizations can be used to create unique user experiences for immersive stereoscopic displays, and present first impressions of the effects as described by observers. However, besides the work by Zhang et al. and Krekhov et al. little work has investigated dichoptic attention cues, and little work has involved testing cue salience through pop-out style tasks [21], [48]. For OST AR displays specifically, we could not find any other work that investigated the saliency of dichoptic visual cues through pop out style tasks. Due to the issues inherent to OST AR displays, it is possible that visual cues that have been shown to have high saliency on other types of displays are less salient on OST AR displays. Therefore, in this work, we revisit the DeadEye cue that was previously explored in VR by Krekhov et al. and investigate several other variations of dichoptic visual cues created by introducing either hue, saturation, or value discrepancies between the observer's eyes [21].

3 EXPERIMENT 1

In this section, we describe an experiment that compares the saliency of dichoptic cues on an OST AR display.

3.1 Participants

We recruited 20 participants ages 18–56 (mean 26.3, SD 9.2) 14 male, 6 female from the population of our university. The participants were screened for exclusion criteria, including pregnancy, history of seizures/epilepsy, neurological and motor impairments, color blindness, strong eye dominance, night blindness, and visual conditions that otherwise impair their visual acuity. Eight participants wore glasses during the experiment and all participants reported having normal visual acuity (with correction if needed).

Participants were asked to rate their level of experience with using stereoscopic displays, such as watching 3D movies or using AR/VR head-mounted displays, using a seven point scale with 1 meaning "least experienced" and 7 meaning "most experienced." Participants reported a mean level of experience of 5.5 with a standard deviation of 1.7. Participants were also asked to rate their level of experience with using monocular displays using the same seven point scale, which resulted in a mean of 2.8 and standard deviation of 1.5.

3.2 Materials

In order to present dichoptic imagery to the participant that differed in manners other than parallax effects, we

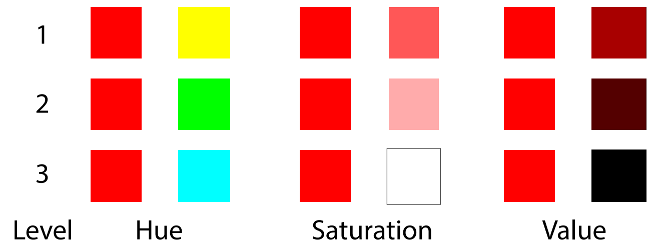


Fig. 1. The figure depicts a side by side view of the visual cues, separated by what each eye would see. Although the original hue is randomized in the experiment, for this illustration the original hue is shown as 0 (red) for the left eye, and is shifted to create a dichoptic cue by changing the color on the right eye. Note that black would appear to be completely transparent on the HoloLens 2.

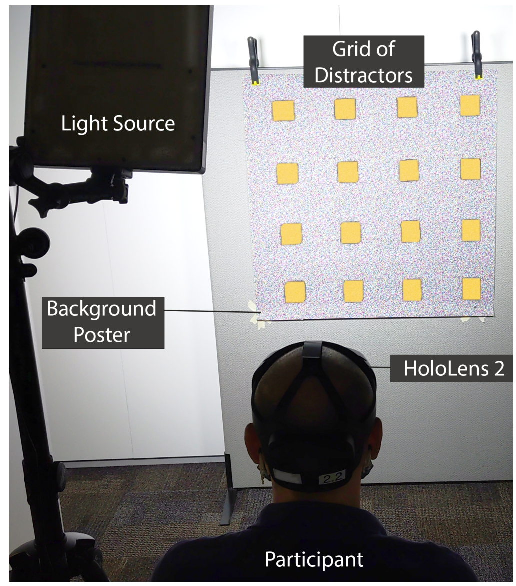


Fig. 2. An annotated depiction of the study setup for experiment 1 taken via the built-in screen capture software of the HoloLens 2. A grid of virtual distractor objects is displayed over a physical background poster. On conditions with a dichoptic visual cue present, the appearance of a randomly-chosen cube will be changed in the imagery presented to one of the participant's eyes.

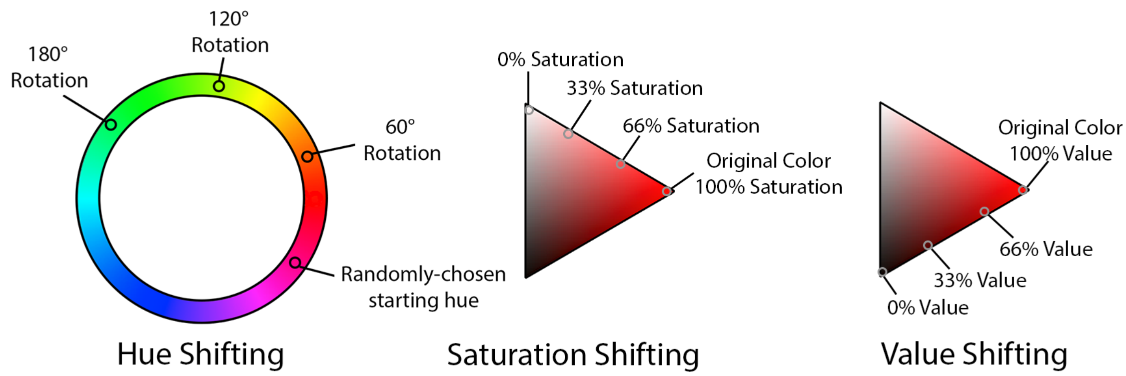


Fig. 3. A depiction of the hue, saturation, and value color shifts used to generate the dichoptic visual cues.

set up a scene in the Unity engine (version 2019.4.26) in which the main camera rig consisted of separate cameras for each of the participant's eyes. Objects in the unity scene were duplicated in place and set to a layer mask such that one object would appear solely to the participant's left eye and one would appear solely to the participant's right eye. In this manner, different materials could be set for each, otherwise identical, object in order to generate binocular rivalry.

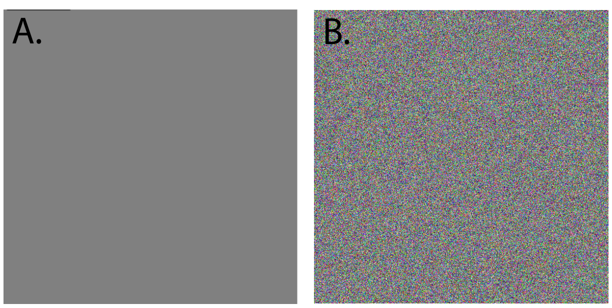


Fig. 4. The figure depicts the two different background images used for the physical background independent variable: the solid grey background (A), and the chromatic aberration background (B).

We investigated three main types of visual cues involving binocular rivalry in this study, each with three different levels of intensity. As the level of intensity increases, the imagery between the participant's eyes differs to a greater degree. As an initial exploration into the effectiveness of these types of visual cues, we investigated the three dimensions of the hue, saturation, value (HSV) color space. This color space is commonly visualized as a cylinder, where hue represents the radial angle, saturation represents the radial length, and value represents the vertical height.

- **Hue:** These cues appeared with a different hue between the participant's eyes (see figure 1 left.) The intensity level was varied by increasing the amount of degrees of rotation away from the randomly chosen initial hue of the object for one of the participant's eyes (see figure 3 left.) The three levels of intensity for this type of visual cue involved 60, 120, and 180 degrees of rotation respectively.
- **Saturation:** These cues appeared with a different saturation level between the participant's eyes (see figure 1 middle), effectively making the image whiter in one eye compared to the other. The intensity level was varied by specifying a percentage of saturation at levels of 0%, 33%, and 66% respectively (see figure 3 middle.)
- **Value:** These cues appeared with a different value level between the participant's eyes (see figure 1 right.) Since an additive light model OST HMD was used to display these cues, where darker colors appear more transparent than lighter colors, reducing the value level has a similar effect to reducing the transparency level of affected object. The intensity level was varied by specifying a percentage of value at levels of 0%, 33%, and 66% respectively (see figure 3 right.) At 0% value levels, the object is completely transparent on the display, making this particular level similar to the DeadEye cue described in section 2.

For each of these types of visual cues, the participant is presented with a different color for each eye, and the relative difference between these two colors can be measured as a distance between two points in the HSV color space. It is possible that this distance correlates with the saliency of the dichoptic cue, and is quantified in terms of the radius, r, and height, h, of the HSV color space. We designed the dichoptic cues in this study to investigate the whole range of the color space available to the AR display. Therefore, the hue-based cues involve HSV color space distances of r, $r\sqrt{3}$, and $2r$ for intensity levels 1, 2, and 3 respectively. The saturation-based cues involve HSV color space distances of $\frac{r}{3}$, $\frac{2r}{3}$, and r for intensity levels 1, 2, and 3 respectively. Finally the value-based cues involve HSV color space distances of $\frac{h}{3}$, $\frac{2h}{3}$, and h for intensity levels 1, 2, and 3 respectively.

It should also be mentioned that due to the additive light model employed by the AR display, the colors of the visual cues observed by participants blend with the color of the physical environment behind it [32]. Since the two colors used in the dichoptic cue are blending towards a single background color, the distance between the two colors of the dichoptic cue will decrease after color blending. The only time this is not the case is if the background is pure black, in which case the distance between the colors will remain unchanged.

Following the example set by Krekhov et al. we arranged our visual cues within a four by four grid of distractor objects [21]. These distractor objects took the form of cubes that were sized to be 1.8 degrees of visual angle with horizontal and vertical spacing between the cubes of 1.2 degrees of visual angle, resulting in a grid that measured 10.8 by 10.8 degrees of visual angle (see figure 2.) These values were somewhat reduced from those used by Krekhov et al. due to the limited vertical field of view of the HoloLens 2. For each trial during the experiment, the hue of the cubes in the grid was set to a randomly chosen number between 0 and 359, while saturation and value were set to 100%. Half of the trials experienced by participants involved a dichoptic cue being applied onto a randomly-chosen cube from the grid, shifting its hue, saturation, or value away from the initial color in one eye, while its appearance remains unchanged in the other eye. The eye chosen to present the shifted color to is determined by the **Presented Eye** independent variable described in section 3.3.1. The remaining half of trials (those without a dichoptic cue) are presented as a homogeneous grid of cubes. Since the visual cues are presented for short durations (250 milliseconds), it is unlikely that participants will experience any alternation in their perception of the appearance of the cue. Therefore their perception of the presented cues is likely influenced by the factors associated with onset binocular rivalry, such as luminance, contrast, and spatial frequency [16].

We chose to use the Microsoft HoloLens 2 as the OST AR display for this study. The HoloLens 2 has a resolution of 2048×1080 per eye, and a field of view of 43×29 degrees[1]. The brightness of the HoloLens was verified to be set to maximum for each participant, and the device was remotely connected to a PC controlling the sequencing of study events via the Unity engine and Holographic Remoting. As with other waveguide based OST AR displays, there are brightness non-uniformities when observing virtual content through the HoloLens 2 [49]. These can be observed monoscopically as variations in the appearance of virtual content within different parts of the field of view of the HoloLens 2, and they can also be observed stereoscopically since these uniformities appear to be in different regions of the field of view for each eye/display. Because of this, and similar to previous works [21], we ensured that the position of the visual cue within the arrangement of distractor objects was randomized for each trial the participant experienced. This randomization of cue position on the display should reduce the overall impact that these non-uniformities have on the results of the study.

We used two background posters during this study, a solid grey background that would introduce a uniform color blending between the AR imagery and the participants' view of their physical environment, and a "chromatic aberration" background that introduced a randomized color blending between the AR imagery and the participants' view of their physical environment (see figure 4.) The posters were printed on premium archival matte paper and measured 0.9x0.9 meters. The solid grey background poster was generated using a uniform grey pixel intensity of 128 on a scale of 0 to 255. The chromatic aberration background poster was generated by randomly assigning pixel values across the range of available color space (randomizing between 0 and 255 for the red, green, and blue color channels independently). The pixels on the chromatic distortion background were sized at 1.58 millimeters, which corresponds to a visual angle of 0.0455 degrees (2.725 arc minutes) at two meters distance. When viewed under the LED lighting from the participant's perspective, the illuminance of the background posters were measured using an Urceri MT-912 light meter prior to starting participant sessions, where the average illuminance was found to be 115 Lux and the standard deviation was 7.65 Lux.

The testing environment was set up so that the participant was seated in a chair two meters in front of a wall, upon which the background posters would be hung. Two LED light sources were positioned just to the left and right sides of the participant, also at a distance of two meters away from the wall, and at a height of 1.83 meters (measured from the top of the light).

1. https://uploadvr.com/hololens-2-field-of-view/

3.3 Methods

3.3.1 Study Design

The experiment consisted of four independent variables:

- **Cue Type: (3)** The type of visual cue presented to the participant, which could consist of color differences in the form of 1. *Hue*, 2. *Saturation*, or 3. *Value*.
- **Cue Intensity: (3)** Each cue had three varying levels of intensity that spanned the range of HSV color space on the HoloLens 2, as described in section 3.2.
- **Physical Background: (2)** There were two physical background posters that were displayed behind the virtual imagery shown on the HoloLens 2, which were 1. *Solid Grey* and 2. *Chromatic Aberration*, as described in section 3.2.
- **Presented Eye: (2)** The color of the visual cue would be changed away from that of the distractor objects on either the participant's 1. *dominant eye* or 2. *non-dominant eye*. For example, if a visual cue was shown within the grid of distractor objects to the participant's dominant eye, then their non-dominant eye observed a grid of homogeneous objects.

We used a 3×3×2×2 within subjects study design where the order of each condition was *randomized* for each factor except the physical background, which was *counterbalanced* to avoid wear and tear on the paper posters. This led to the experiment taking place in two segments, the first of which consisted of all conditions for one particular background poster, after which the poster was exchanged and the remaining conditions were presented. Each condition was presented to the participant in a block of ten trials, which were randomized so that half of them had a visual cue within the grid of distractor objects and the other half did not have a visual cue and consisted of homogeneous distractor objects. These ten repeated trials were all performed in succession, which allowed for calculation of measures such as false negatives and false positives for each condition.

3.3.2 Measures

The main objective measure of the experiment involves the accuracy of the participants' responses, and can be broken down into several combinations based on whether or not a visual cue was present for a particular trial and whether the participant observed the presence of the visual cue:

1) Correct Response: The participant correctly indicated the presence or absence of the visual cue.
2) False Positive: The visual cue was not presented within the grid and the participant responded that they observed the presence of the stimulus.
3) False Negative: The visual cue was presented within the grid and the participant responded that they did not observe the presence of the stimulus.

Based on these possible combinations, we examined error rate, which was the total number of false positives and false negatives divided by the total number of trials (10) for each condition. We also individually examined the false negative rate and false positive rate for each condition.

3.3.3 Procedure

As participants arrived to the testing environment, they were asked to read an informed consent document and provide their verbal consent to participate in the study. While the participant read the consent form, the HoloLens 2 was sanitized via a UV box along with all other equipment the participant would come into direct contact with. Participants were then asked to perform a dominant eye test by performing the hole-in-card test at a distance of two meters to a target on a wall directly in front of them. Participants were then asked to sit in a chair positioned two meters in front of the physical background poster that would be used for their first section of study conditions. The experimenter then explained how to properly don/doff the HMD, and asked the participant to then put on the HMD and hold a keyboard on their lap. Following this, the experimenter opened the virtual scene on the HMD and explained the instructions and sequencing of the study procedure.

For each condition, participants were shown a message and two virtual cursors within the HMD. The message directed the participant to align the two AR virtual cursors while keeping their gaze fixated on the cursors, one of which was world-fixed to the center of the physical background poster positioned two meters in front of the participant, and the other was head-fixed to the participant, also at a distance of two meters. Once participants had aligned the two cursors, they pressed the space button to begin the trial. After pressing space, the message disappeared and only the two cursors were visible for a duration of 2.5 seconds. Following this, the cursors disappeared and the 4×4 grid of cubes appeared for a duration of 250 milliseconds and then disappeared. The participant was then asked whether or not they noticed that one of the cubes appeared differently than the others. The participant pressed 'y' or 'n' on the keyboard for "yes" or "no", and were then returned to the initial screen with the two cursors to begin the next trial.

After ten repeated trials of one condition (five with a cue present, five without), the next condition was randomly chosen from a list of remaining conditions. This process repeated until all conditions for the first background poster were completed. Following this, the experimenter exchanged the background posters, and the process repeated again for all conditions involving that poster.

Upon completing all conditions, the participant was instructed to remove the HMD and place it and the keyboard on a table next to them. Finally, the participant completed a demographics questionnaire and was compensated with fifteen dollars for their time. This procedure took between 45–55 minutes per participant depending on how quickly the participants moved through the conditions.

3.3.4 Hypotheses

Based on the previous literature, we formulated the following hypotheses:

- **H1** Hue-based visual cues will be more effective than saturation and value-based cues.
- **H2** Visual cues with higher intensity levels will be more effective (lower error rates) compared to cues with low intensity levels.
- **H3** Visual cues will have similar effectiveness (similar error rates) when the shifted color of the dichoptic cue is presented to the participant's dominant eye compared to when it is presented to their non-dominant eye.
- **H4** Visual cues will be more effective (lower error rates) when displayed over uniform backgrounds compared to visually complex backgrounds.

4 RESULTS 1

This section describes the results of the first experiment gathered through analysis of the participants' error rate, number of false negatives, and number of false positives. All analysis took place using SPSS version 28.0.0.0.

Since each condition was repeated in ten trials, prior to analysis these ten trials were aggregated via SPSS to generate mean error rate, mean false negatives, and mean false positives for each condition. Following this, the aggregated data was analyzed with a repeated-measures ANOVA with four factors (cue type, cue intensity, physical background, and presented eye) with levels of 3, 3, 2, and 2 respectively. From this, Tukey multiple comparisons tests were performed with Bonferroni correction at the 5% significance level. We confirmed the normality of the results using Shapiro-Wilk tests set to 5% level and QQ plots.

Our results indicated several significant main effects ($p<0.05$). Cue type was found to have a significant effect on mean error rate, F(2,38)=106.91, $p<0.001$, $\eta_p^2 = 0.85$, and mean false negatives, F(2,38)=118.38, $p<0.001$, $\eta_p^2 = 0.86$. No significant effect was found on the number of false positives, F(2,38)=1.02, p=0.37, $\eta_p^2 = 0.05$. Pairwise comparisons revealed significant differences both for error rate and false negatives for each comparison of cue types ($p<0.001$). These results indicated that value-based cues had the highest error rate (m=0.462, SE=0.011) and false negatives (m=0.399, SE=0.018). Hue-based cues had the lowest error rate (m=0.209, SE=0.020) and false negatives (m=0.157, SE=0.017). Finally, saturation-based cues fell between them in both error rate (m=0.361, SE=0.018) and false negatives (m=0.298, SE=0.015). Figure 6 shows a comparison of the false negative rates broken down by the three cue and the three cue levels.

Cue intensity level was found to have a significant effect on mean error rate, F(2,38)=42.93, $p<0.001$, $\eta_p^2 = 0.69$, and false negatives, F(2,38)=46.77, $p<0.001$, $\eta_p^2 = 0.71$. No significant effect was found for false positives, F(2,38)=0.82,

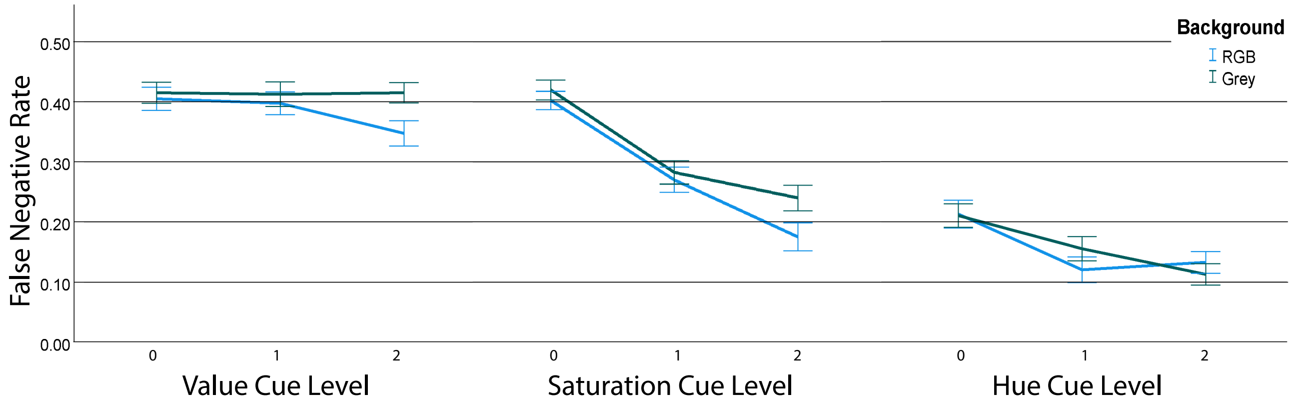


Fig. 5. This figure depicts a comparison of the false negatives from experiment one across cue type, cue level, and background. Error bars represent standard error with a multiplier of 1.

p=0.45, $\eta_p^2 = 0.04$. Pairwise comparison revealed significant differences for both error rate and false negatives for each comparison of cue types (all p<0.01). These results indicate that overall error rate and false negative error rate reduce with increased cue intensity level, and therefore with increased amounts of color distance between colors shown between the participants' eyes. Cue level 0, which involved the least amount of color distance between the participants' eyes, yielded the highest error rate (m=0.406, SE=0.011) and the highest false negatives (m=0.344, SE=0.014). Cue level 1 yielded lower error rate (m=0.335, SE=0.017) and false negatives (m=0.273, SE=0.016). Finally, cue level 2, which involved the greatest color distance between the participants' eyes, yielded the lowest error rates (m=0.292, SE=0.018) and false negatives (m=0.237, SE=0.017). Again, this effect can be seen in figure 6, as well as figure 5, which breaks down the results further by including cue type, cue level and background.

Physical background was found to have a significant effect on mean error rate, F(1,19)=10.79, p=0.004, $\eta_p^2 = 0.36$, and mean false negatives, F(1,19)=5.47, p=0.03, $\eta_p^2 = 0.22$. No significant effect was found on the number of false positives, F(1,19)=0.29, p=0.59, $\eta_p^2 = 0.02$. This effect on false negative rates can be seen in figure 5.

Finally, presented eye was not found to have any significant effects on error rate, F(1,19)=0.33, p=0.58, $\eta_p^2 = 0.02$, false negatives, F(1,19)=0.44, p=0.52, $\eta_p^2 = 0.02$, or on false positives, F(1,19)=0.01, p=0.93, $\eta_p^2 < 0.01$.

When examining the interaction effects, several significant results were found. Here we report only the significant interaction effects. There was a significant interaction effect between cue type and cue level on error rate, F(4,76)=12.68, p<0.001, $\eta_p^2 = 0.40$, and on false negatives, F(4,76)=15.38, p<0.001, $\eta_p^2 = 0.45$. There was a significant interaction effect between physical background, cue type, and cue level on false negatives, F(4,76)=2.79, p=0.03, $\eta_p^2 = 0.13$. There was a significant interaction effect between physical background, cue level, and the presented eye on false negatives, F(2,38)=6.42, p=0.004, $\eta_p^2 = 0.25$. Finally, there was a significant interaction effect between all four independent variables on both error rate, F(4,76)=3.28, p=0.016, $\eta_p^2 = 0.15$, and on false negatives, F(4,76)=3.18, p=0.018, $\eta_p^2 = 0.14$.

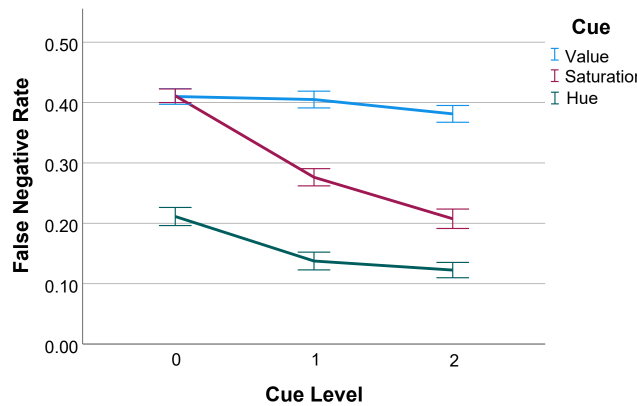


Fig. 6. This figure depicts a comparison of the false negatives from experiment one across cue type and cue level. Error bars represent standard error with a multiplier of 1.

4.1 Hue

We performed an additional analysis on the overall effects of the hue of the objects within the grid of stimuli on the number of false negatives experienced by the participants. Since hue was randomized for each individual trial of each condition, our results showed a relatively uniform distribution of trials across all hues in the range of (0–359) for the HSV color space. Assuming that the hue of the visual cue had no impacts on the rate of false negatives, then we should expect to see a uniform distribution in the plots of hue versus false negative rate, with a horizontal linear fit line. However, as shown in figure 7, we see that this is not the case, and the false negative rate increases for hue values near 120, which correspond to green-tinted hues. For reference, red corresponds to both 0 degrees and 360 degrees of hue rotation, green corresponds to 120 degrees, and blue corresponds to 240 degrees. Page 1

We examined this in SPSS via use of the locally estimated scatterplot smoothing (Loess) fit line (shown in green in the figure). From this fit line, we observed a distribution with its mean centered at 125.8 hue on the x-axis with a standard deviation of approximately 60. The amplitude was measured to be approximately 0.14, and the vertical shift was measured to be 0.23.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3195111

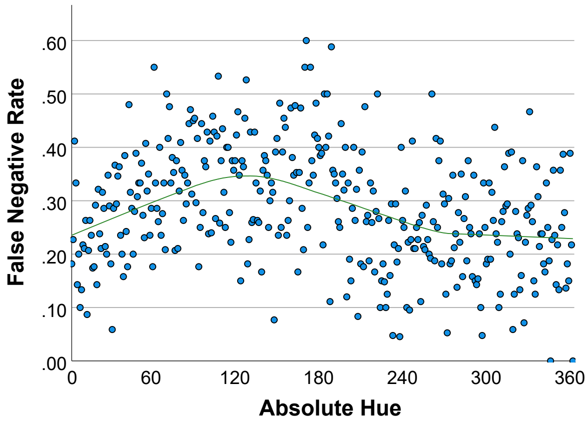


Fig. 7. This figure depicts the overall effects of the hue of the grid of stimulus objects on the false negative rate from experiment one. The green line represents a locally estimated scatterplot smoothing (Loess) fit set to include 50% of points in the data set and using an Epanechnikov kernel.

5 Experiment 2

In this section, we describe an experiment that investigates the subjective qualities of the visual cues investigated in the previous experiment. All participants from the first experiment participated in this experiment immediately after completing the first, therefore all participant data is the same and can be seen in section 3.1. Our experimental setup was also exactly the same as the previous setup and is described in section 3.2.

5.1 Methods

5.1.1 Study Design

The experiment consisted of two independent variables:

- **Cue Type: (3)** The type of visual cue presented to the participant, which could consist of color differences in the form of 1. *Hue*, 2. *Saturation*, or 3. *Value*.
- **Cue Intensity: (3)** Each cue had three varying levels of intensity that spanned the range of color space on the HoloLens 2. These intensities were set to the same values used in experiment one.

We used a 3×3 within subjects study design where the order of each condition was randomized for each factor. For this study, all conditions were presented solely in front of the solid grey background poster. The "chromatic aberration" poster was not used because we wanted the participants to be able to observe the visual cues with a uniform color blending between the virtual imagery and the physical environment. Each condition was presented to the participant along with a user interface depicting three different questions. Participants would respond to these three questions using the number keys on the keyboard, and then the next condition would be displayed. In this section, no distractor objects were used, and participants were free to observe the conditions without time constraints while responding to the questions.

5.1.2 Measures

The measures for experiment two consisted of three subjective prompts for the participants to respond to.

1) **Noticeability**: Participants were asked to rate how noticeable the visual cue was compared to the other cues they observed. They rated noticeability using a 9-point scale with 1 meaning least noticeable and 9 meaning most noticeable.
2) **Urgency**: Participants were asked to rate the sense of implied urgency of the visual cue compared to the other cues they observed. "Urgency" was described using an anecdote in which a visual cue was being used as a system notification, and could be a low-urgency notification such as a new text message, or a high urgency notification such as an amber/disaster alert. Participants responded to the prompt using a 9-point scale where 1 meant least urgent and 9 meant most urgent.
3) **Comfortability**: Participants were asked to rate the level of general visual comfort experienced while observing the visual cue while responding to the three subjective prompts. Participants used a 9-point scale where 1 meant least comfortable and 9 meant most comfortable.

5.1.3 Procedure

Upon finishing experiment one (described in section 3), the participants immediately started experiment two. Participants were instructed to look to the center of the solid grey background poster presented in front of them at a distance of two meters. The experimenter then started the Unity scene for experiment two on a remote PC via Holographic Remoting.

Upon loading, the participant would see a single randomly chosen visual cue from the set of conditions for experiment two. Above the cue was a user interface that initially asked the participant to rate the level of noticeability of the cue using a scale with a range of one to nine. The participant indicated their response by pressing the number on the keyboard, after which the user interface would update to ask the participant to rate the level of implied urgency of the visual cue. The participant would indicate their response in the same manner, then the user interface would update to ask the participant to rate the level of comfortability of the visual cue. Once the participant responded to this third prompt, a new randomly-chosen cue was picked from the list of remaining visual cues, and the process repeated. Once the participant had responded to all nine visual cues, they were instructed to remove the HoloLens 2.

5.1.4 Hypotheses

Based on the results of the experiment one and those found in the literature, we formulated the following hypotheses:

- **H5** Hue-based visual cues will be rated as more noticeable, more urgent, and less comfortable compared to saturation-based cues and value-based cues.
- **H6** Dichoptic visual cues displayed at higher intensity levels will be rated as more noticeable, more urgent, and less comfortable compared to lower intensity levels.

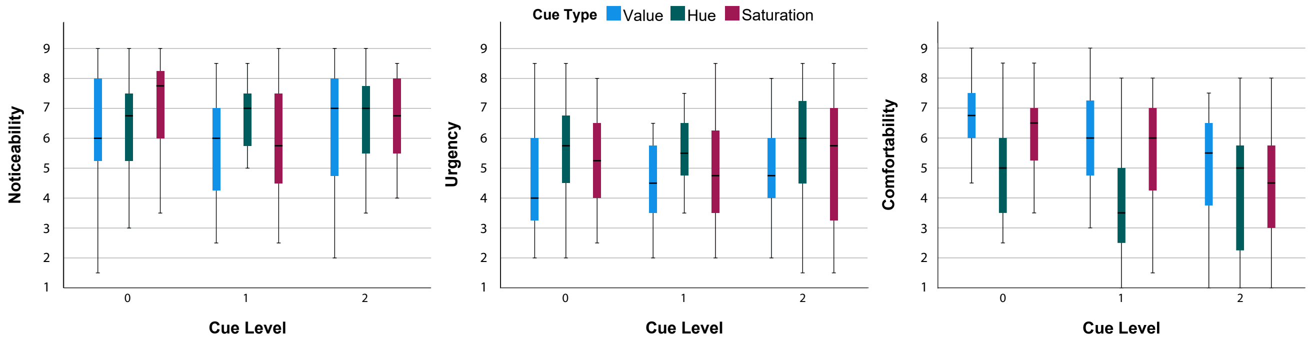


Fig. 8. This figure depicts the subjective noticeability, urgency, and comfortability results from experiment 2.

6 Results 2

This section describes the results of experiment two. All analysis took place using SPSS version 28.0.0.0. Similar to experiment one, the data was analyzed with a repeated-measures ANOVA with two factors (cue type, cue intensity) with levels of three and three respectively. From this, Tukey multiple comparisons tests were performed with Bonferroni correction at the 5% significance level. We confirmed the normality of the results using Shapiro-Wilk tests set to 5% level and QQ plots.

Our main results can be seen in figure 8. Our results indicated several significant main effects (p<0.05). Cue type had a significant main effect on the sense of implied urgency of the cue, F(2,38)=3.85, p=0.03, $\eta_p^2 = 0.17$, and on the comfortability of the cue, F(2,38)=11.82, p<0.001, $\eta_p^2 = 0.38$. No significant main effect was found for cue type on the noticeability of the cue, F(2,38)=2.17, p=0.13, $\eta_p^2 = 0.10$. Pairwise comparisons revealed a significant difference (p=0.023) in the sense of implied urgency between the value-based cues (m=4.667, SE=0.252) and the hue-based cue (m=5.592, SE=0.278), indicating significantly higher sense of urgency for the hue-based cues. They also revealed a significant difference (p<0.001) in comfortability between the hue-based cues (m=4.450, SE=0.386) and the value-based cues (m=5.942, SE=0.254), indicating that value-based cues were more comfortable than the hue-based cues. A second significant difference in comfortability (p=0.029) was found between the hue-based cues and the saturation-based cues (m=5.308, SE=0.312), indicating significantly higher comfort in the saturation-based cues compared to the hue-based cues.

Cue level was found to have a significant main effect on the comfortability of the cue, F(2,38)=9.51, p<0.001, $\eta_p^2 = 0.33$. No significant effects were found for cue level on noticeability, F(2,38)=1.52, p=0.23, $\eta_p^2 = 0.07$, or on the sense of implied urgency, F(2,38)=1.48, p=0.24, $\eta_p^2 = 0.07$.

Pairwise comparisons revealed a significant difference in comfort (p=0.008) between cue intensity level 0 (m=5.975, SE=0.239) and level 1 (m=5.108, SE=0.313), as well as a significant difference (p=0.006) between cue intensity level 0 and level 2 (m=4.617, SE=0.401). No significant difference was found between cue intensity level 1 and level 2.

7 Discussion

As presented in the 2018 survey by Kamkar et al. the saliency of a visual cue is determined by many different factors, including the cue's appearance in relation to that of distractor objects and the background, as well as factors such as binocular rivalry [6]. In the following sections, we show that the salience of some of the investigated cues can be modeled as a function of HSV color space distances between the two colors that comprise dichoptic cues. These color space distances were previously established in section 3.2 and will be discussed in detail in the proceeding sections.

7.1 Saliency of Color-Based Dichoptic Cues

Through experiment one, hue-based visual cues were found to have the lowest error rates in the pop-out task, followed by saturation-based cues and then finally value-based cues. Further, hue-based visual cues were rated as appearing significantly more urgent compared to the value-based cues. From this, we accept hypothesis **H1** and partially accept hypothesis **H5**. Intensity level was also shown in experiment one to have significant impacts on the error rates of participants, where increased intensity level led to reduced error rates, which leads us to accept hypothesis **H2**. However, in experiment two, intensity level was only shown to have effects on the perceived comfort of the dichoptic cue and no effects on noticeability or urgency were found, therefore we can only partially accept hypothesis **H6**.

As mentioned above, it is possible that the HSV color space distances introduced in section 3.2 can be used to predict the saliency of the dichoptic cues. In comparing between the hue-based cues and the saturation-based cues, they both involve HSV color space distances that rely on the radius of the cylindrical color space. Further, they both involve an intensity level with the same distance, where the hue-based cue at intensity level 0 and the saturation-based cue at intensity level 2 both have a distance equivalent to r, the radius of the cylindrical color space. If the HSV color space distance between the two colors used in the dichoptic cue is an effective way to compare saliency levels of the dichoptic cues, then we would expect to see overlapping values for these two specific conditions. Looking at figure 6, we can see that this is the case, and that the false negative values for these two conditions do indeed overlap and fall within the range of a standard error (multiplier 1) of each other. However, because we did not originally plan to

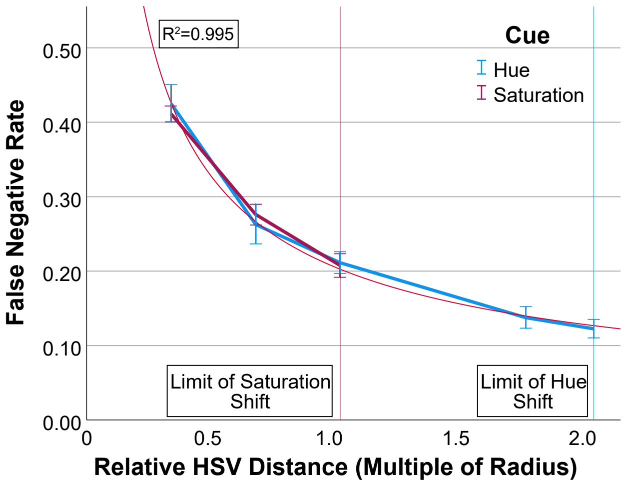


Fig. 9. This figure compares the false negative rates between the hue and saturation based cues in terms of relative HSV color space distance between the two colors shown to each eye. It is measured in multiples of the radius of the color space. This figure includes data from four additional participants that solely evaluated the HSV distances of $\frac{r}{3}$, and $\frac{2r}{3}$ for the hue-based cues to create the points of overlap between the hue and saturation cues. The equation for the fit line is in the form of $y = ax^b$, where $a = 0.203$ and $b = -0.676$, which has an $R^2 = 0.995$. The vertical lines represent the maximum color space distances that can be achieved between two colors via a saturation-based cue (x=1r) or a hue-based cue (x=2r) alone.

compare the study conditions in this manner, we only have one point of overlap in the data for these two cue types, and we do not know whether or not hue-based cues would follow the same pattern as the saturation-based cues when measured at the color space distances of $\frac{r}{3}$, and $\frac{2r}{3}$.

For the purpose of this discussion, we ran four additional participants through the study procedure of experiment one, testing only hue-based cues at HSV color space distances of $\frac{r}{3}$ and $\frac{2r}{3}$, so that these conditions could be compared with the saturation-based conditions at the same HSV distances. The data from these participants was plotted along with the initial data from experiment one, and can be seen in figure 9. It appears as though participants perform similarly with hue-based cues and saturation-based cues, and that the differences measured previously were a factor of color space distance and not necessary cue type. This implies that HSV color space distance is an effective way to compare the saliency level of color-based dichoptic cues.

We performed a curve estimation regression on the data from figure 9 via SPSS that revealed that a power model best fits the data. This curve can be seen as a red line in the figure, and the equation and R^2 values for this model are reported in the caption of the figure. While we acknowledge that four participants is not necessarily an adequate sample size for confirming this, we include the data here as a method of explaining the results of the above studies, and emphasize that this should be evaluated further in future work.

With this model, we can see that hue-based dichoptic cues achieve lower error rates compared to saturation-based cues because they are able to achieve higher HSV color space distances, and thus higher levels of saliency, compared to the saturation-based cues. Saturation-based cues are limited in that the maximum color space distance they can achieve (without combination with other cues) is equal to the radius of the cylindrical HSV color space, since 0 saturation refers to a point on the central vertical axis of the cylinder and 100% saturation refers to a point on the outer edge. Hue-based dichoptic cues have a maximum distance of two times the radius of the color space, which is achieved by performing a 180 degree hue rotation at 100% saturation within the color space.

Unfortunately, it is more difficult to compare value-based cues in a similar manner because value-based cues have a maximum distance determined by the height of the color space, and we have not established a relationship between the radius and height of the color space. However, the intensity levels of the value-based cues had HSV color space distances with the same multiples of height that the saturation-based cues had as multiples of the radius. Therefore, it is likely that the height of the color space must be less than the radius, which would mean that value-based cues would be less salient than the saturation-based cues, which is supported by our data.

While this is certainly a strong possibility, it is not something that we can determine from this work alone. In order to further evaluate the relationship between value-based dichoptic cues and the other cue types, future work should investigate value-based dichoptic cues in combination with hue or saturation-based cues. For example, comparing user performance in conditions with a variable intensity hue-based cue alone to conditions consisting of both a variable hue and variable value shift. From this, a potential relationship between the radius and height of the color space can be determined and then value-based cues could be directly compared with hue or saturation-based cues in the saliency model established above.

Another potential reason for the increased false negative rates in the value-based conditions is due to the brightness non-uniformities of the HoloLens 2. Several of our participants mentioned that the appearance of the imagery shown on the display changed based on where it was positioned within the field of view of the device. While the cursor alignment portion of the procedure ensured that the grid of objects would appear in the same portion of the device's field of view for each trial, participants were able to notice this effect on the text-based user interfaces when orienting their heads downward to look at the keyboard and then back forward towards the background poster. This is also something that we noticed when implementing the project and observing the grid of distractor objects for durations longer than the 250 milliseconds it was displayed for during the study. It is possible that this effect in the grid of objects would make some of the objects appear similar to some of the value-based cues, particularly those displayed at the lower intensity levels. Since this effect should be equally apparent for each trial of the study, it makes sense that participants would treat this effect as a part of the regular appearance of the virtual imagery, and therefore make false negative errors. Future work could determine if this is the case by seeing if user performance improves in a similar experiment using a different OST AR display with less brightness non-uniformities, such as a beam-splitter based OST AR display. This could also potentially be measured on the HoloLens 2 by performing a task in which users observe cues in randomized positions within the device's FOV and are tasked with indicating whether or not the cue is a value-based dichoptic cue or a normal stereoscopic cue.

7.2 Effect of Eye Dominance

As mentioned in section 2, the literature was inconclusive as to whether or not eye dominance affects user perception of dichoptic imagery. Do to the similarity of our methods with the work of Krekhov et al. and Logothetis et al. in which the dichoptic imagery is only presented to the user for fractions of a second, we hypothesized in **H3** that there would be a similar effectiveness (similar error rates) between conditions in which the shifted color of the dichoptic cue was presented to the participant's dominant eye versus their non-dominant eye [3], [50]. We observed similar magnitudes of error rates in these conditions in line with this hypothesis. This result is convenient for the potential applications of dichoptic cues in AR displays, as it implies it may not be necessary for the designer of the AR imagery to know the dominant eye of the participant in order to achieve an effective dichoptic cue. Instead, the designer can produce a color shift away from the virtual cues' original color in either eye, while keeping the image the same on the other eye. While all of the conditions investigated in this paper used this technique of creating a dichoptic cue, it is possible that different techniques could be used to gain a similar, or perhaps even better, effect. For example, a slightly different approach could be taken where dichoptic cues are generated via color shifts that are performed for both of the participant's eyes in opposite directions across the color space. In this manner, the cue would likely stand out more from homogeneous distractor objects of the original color in a pop-out task, since neither eye is observing a homogeneous grid, however it is less intuitive to imagine how cues generated by these two techniques would compare in a more practical task, such as a timed search task where participants respond to dichoptic cues positioned within their physical-virtual environment.

Since saliency was only measured in this work when participants were observing dichoptic imagery for brief durations of 250 milliseconds at a time, it remains unclear as to whether effects of eye dominance would come into play should the dichoptic cues be presented for longer durations of seconds/minutes at a time. For longer durations, it may be the case that the imagery from the observer's non dominant eye is suppressed more often than the imagery from the dominant eye, which means that a dichoptic visual cue could potentially be more salient when presented to the dominant eye. However, as we will discuss later in sections 7.4 and 7.5, the results of experiment 2 suggest these cues tend to be uncomfortable when presented for longer durations, so using similar dichoptic cues in this manner may not be advisable.

7.3 Effects of Background Appearance

An interesting result of experiment one is that participants made significantly fewer errors when viewing the cues over the chromatic aberration pixelated background compared to the solid grey background. We hypothesized in **H4** that this would be the other way around, therefore we cannot accept this hypothesis.

In forming hypothesis **H4**, we thought that cues presented in front of the uniform background would be easier to identify since there is a uniform color shifting of the cue when the imagery from the AR display and the environment blends together. However, due to color blending, this background solely introduces a luminance shift when the virtual image of the cue blends with that of the physical background. On the other hand, the chromatic aberration background is comprised of pixels with different hues and luminances. So in this case, when the colors of the physical background and virtual dichoptic cue mix, the resulting blended color has both a luminance shift and a hue shift away from its original color. This background also has a higher spatial frequency compared to the uniform background.

The literature on saliency of visual cues has previously established that saliency is increased when multiple features differ between the cue and the distractor objects [6], [51], [52], therefore its possible that the presence of hue-based features and/or spatial frequency-based features in addition to the luminance-based feature contributed to the higher saliency of visual cues displayed over the chromatic distortion background poster.

Turatto and Galfano investigated hue and luminance independently in 2000, where they found that either of these features can be an effective way of capturing the attention of the observer [53]. In 2009, Engmann et al. similarly found that both luminance and hue based gradients applied over a scene significantly bias points of fixation made by their participants, and that when both types of gradients were applied simultaneously their effects combined linearly to produce a stronger bias in participant fixations [54]. Therefore, it is possible that the hue and luminance features introduced through color blending on the chromatic aberration background conditions combined in a similar manner to produce improved saliency in the AR dichoptic cue.

This result is promising when considering its implications for AR dichoptic cues "in the wild," since many of the environments that users of pervasive AR systems will find themselves in will contain a multitude of different hues and levels of visual complexity in the physical environment. This implies that such colorful or "noisy" scenes would actually be beneficial in that AR dichoptic cues presented in them are actually easier to detect compared to the more uniform and controlled laboratory settings where these studies typically take place. In such environments, it is also possible that the AR dichoptic cues could combine with the relative motion of the user or of elements in their environment to produce an additional feature that further increases the saliency of the AR cues. Such effects should be investigated in future work.

7.4 Subjective Qualities of Dichoptic AR Visual Cues

In experiment two, we hypothesized in **H5** that hue-based dichoptic cues would be rated as more noticeable, more urgent, and less comfortable to observe compared to the saturation-based and value-based cues. Part of this hypothesis was motivated by the results of experiment one, where hue-based cues significantly out-performed the other types of cues, so we expected this would carry over in terms of subjective noticeability and urgency. The comfort portion of this hypothesis was motivated by the observation that the hue-based cues involved larger color space distances between the colors shown to each of the participant's eyes. We did not find any significant effects in terms of noticeability of the cues in experiment two, and found that hue-based cues were only more urgent compared to value-based cues and not saturation-based cues. However, hue-based cues were less comfortable compared to both of the other types of cues, therefore we partially accept **H5**.

We also hypothesized in **H6** that as the intensity level of the dichoptic cues increases, their noticeability and sense of implied urgency would increase while the sense of comfort would decrease. The results of experiment two revealed that the intensity level of the cues only had significant effects on the comfortability of the cues, and that as intensity level increased, comfortability decreased, therefore we also partially accept **H6**.

These effects involving the comfortability of the dichoptic cues present an interesting challenge in using the cues effectively "in the wild," as the cues that were most effective at being noticed by participants were also rated as the most uncomfortable to observe. We should note that in experiment two, participants were allowed to examine the dichoptic cues for as long as they wanted to while indicating their responses on the scales, whereas in experiment one, participants were only exposed to the presence of the cues for durations of 250 milliseconds at a time. When integrating these types of cues into the applications that will use them, exposure time should therefore be carefully considered. Applications could make use of eye-tracking in order to determine that the user has shifted their attention to the cue, after which the cue's appearance could be shifted back to its original (non-dichoptic besides parallax effects) form. In this manner, potentially negative comfortability effects could at least be reduced and managed.

In terms of the sense of implied urgency of the cues, the only significant difference observed was between the hue-based cues and the value-based cues, which are respectively the most and least salient cues according to the results of experiment one. It makes sense that the most salient cues are perceived as implying a more urgent message or notification, while the least salient cues are perceived as implying one that is less urgent or more casual. If these types of visual cues are used in future applications, then these perceptions should be considered when pairing a cue with a message or notification, and perhaps hue-based dichoptic cues should be reserved for cases where the message or notification is of high urgency or importance.

It is interesting that no effects were found in experiment two in terms of the subjective noticeability of the different dichoptic cues. We believe that this may be due to the way in which the cues were presented to participants in this experiment. In experiment two, no grid of distractor objects were employed, and the cues were presented one by one. Therefore, it may have been difficult for participants to accurately rate how noticeable each of the cues were, since there was nothing else in the virtual scene to compare it with. This could potentially be revisited in future work, by presenting pairs of cues and comparing subjective factors such as noticeability or implied urgency between them.

7.5 Using Dichoptic Cues on OST AR Displays

In general, we envision dichoptic AR cues being used to draw the attention of the user to a visual feature within their field of view. A requirement for using dichoptic cues in this manner is that the system needs to know which region of the physical-virtual scene requires the user's attention, then the dichoptic cue can be created by temporarily changing the color of the particular region for one (or both) of the user's eyes over the course of a short duration (e.g. 250 milliseconds). However, there may be situations where the system is unaware of the user's intent, and therefore does not know where to direct the attention of the user. This particular type of situation was investigated previously by Orlosky et al. where they compared different visualization modes that combine visual and thermal information [47]. In their work, the user is tasked with identifying a target object within the scene, where the different visualization modes create differences in saliency between objects in the scene based on their temperature. They found that for this particular type of task, users performed the best with a visualization mode consisting of a static noise effect, where the intensity of the static effect was used to convey temperature. This static visualization significantly outperformed a dichoptic visualization mode, where the intensity of luster/saturation-based dichoptic imagery indicated the temperature of objects in the scene. From their results, it suggests that the differences in saliency that are introduced when using the static noise style visualization are greater than those introduced when using the dichoptic visualization. Therefore, for situations where the system is unaware of a particular region to draw the user's attention to, dichoptic visualizations may not be as well-suited as other visualization styles for increasing saliency differences in the scene.

For tasks where the system knows a particular region of the scene that requires the user's attention, we recommend using hue-based dichoptic cues over saturation-based or value-based dichoptic cues. These cues have the widest available range of color space to work with compared to the other two cue types, and thus have the most versatility when it comes to creating cue types at different intensity levels. It is possible that this type of cue may be a robust way to draw the attention of the user, since it does not necessarily rely on understanding the appearance and structure of the user's physical-virtual environment in order to increase saliency. Future work should examine how hue-based dichoptic cues compare with other non-dichoptic cues in at least a subset of many varied environments that future pervasive AR users will find themselves in.

With regard to the recommended color of dichoptic visual cues, experiment one revealed the interesting result that green-hued dichoptic cues performed worse compared to cues that were more red or more blue (see figure 7.) This finding is similar to what was shown by Etchebehere and Fedorovskaya in 2017, where they noted a similar decrease in noticeability of visual cues for this color region [55]. As a result of this, we recommend avoiding green-hued dichoptic cues in applications where high saliency of the visual cue is important.

8 CONCLUSION

In this paper, we demonstrated that hue-based dichoptic visual cues may be effective as preattentive cues on OST AR displays. In particular we found that users have better performance in a pop-out style task when using hue-based dichoptic cues compared to saturation-based or value-based cues. However, we also found that user comfort decreases as the effectiveness of the cues increases, presenting an interesting compromise in their potential use cases.

Moving forward, dichoptic AR cues should be evaluated against other traditional cue types to establish whether there are other benefits or drawbacks to their usage. Such studies should continue to carefully consider the effects that color blending has on the effectiveness of the AR cues, and how an effective cue can be chosen for the user's dynamic physical-virtual environment without the need for computationally-expensive real-time analysis of scene content.

ACKNOWLEDGMENTS

This material includes work supported in part by the National Science Foundation under Award Number 1564065 (Dr. Ephraim P. Glinert, IIS) and Collaborative Award Numbers 1800961, 1800947, and 1800922 (Dr. Ephraim P. Glinert, IIS) to the University of Central Florida, University of Florida, and Stanford University respectively; the Office of Naval Research under Award Number N00014-21-1-2578 (Dr. Peter Squire, Code 34); and the Advent Health Endowed Chair in Healthcare Simulation (Prof. Welch).

REFERENCES

[1] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht, "Towards pervasive augmented reality: Context-awareness in augmented reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 6, pp. 1706–1724, 2017.

[2] J. M. Wolfe and S. L. Franzel, "Binocularity and visual search," *Perception & Psychophysics*, vol. 44, no. 1, pp. 81–93, Jan. 1988.

[3] A. Krekhov and J. Krüger, "Deadeye: A Novel Preattentive Visualization Technique Based on Dichoptic Presentation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 936–945, Jan. 2019, conference Name: IEEE Transactions on Visualization and Computer Graphics.

[4] A. Erickson, K. Kim, G. Bruder, and G. F. Welch, "Exploring the limitations of environment lighting on optical see-through head-mounted displays," in *Symposium on Spatial User Interaction*, ser. SUI '20. New York, NY, USA: Association for Computing Machinery, 2020.

[5] J. L. Gabbard, M. Smith, C. Merenda, G. Burnett, and D. R. Large, "A perceptual color-matching method for examining color blending in augmented reality head-up display graphics," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2020.

[6] S. Kamkar, H. A. Moghaddam, and R. Lashgari, "Early visual processing of feature saliency tasks: A review of psychophysical experiments," *Frontiers in Systems Neuroscience*, vol. 12, p. 54, 2018.

[7] H.-C. Nothdurft, "Feature analysis and the role of similarity in preattentive vision," *Perception & psychophysics*, vol. 52, no. 4, pp. 355–375, 1992.

[8] C. Ware, *Information visualization: perception for design*. Morgan Kaufmann, 2019.

[9] H.-C. Nothdurft, "Salience from feature contrast: additivity across dimensions," *Vision research*, vol. 40, no. 10-12, pp. 1183–1201, 2000.

[10] L. Huang and H. Pashler, "Quantifying object salience by equating distractor effects," *Vision Research*, vol. 45, no. 14, pp. 1909–1920, 2005.

[11] A. Treisman and G. A. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, pp. 97–136, 1980.

[12] B. McElree and M. Carrasco, "The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 25, no. 6, p. 1517, 1999.

[13] J. Shen, E. M. Reingold, and M. Pomplun, "Guidance of eye movements during conjunctive visual search: the distractor-ratio effect." *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, vol. 57, no. 2, p. 76, 2003.

[14] R. Blake, "A primer on binocular rivalry, including current controversies," *Brain and mind*, vol. 2, no. 1, pp. 5–38, 2001.

[15] D. Roumani and K. Moutoussis, "Binocular rivalry alternations and their relation to visual adaptation," *Frontiers in Human Neuroscience*, vol. 6, p. 35, 2012.

[16] J. Stanley, J. D. Forte, P. Cavanagh, and O. Carter, "Onset rivalry: The initial dominance phase is independent of ongoing perceptual alternations," *Frontiers in Human Neuroscience*, vol. 5, p. 140, 2011.

[17] H. D. Crane and T. P. Piantanida, "On seeing reddish green and yellowish blue," *Science*, vol. 221, no. 4615, pp. 1078–1080, 1983.

[18] V. A. Billock, G. A. Gleason, and B. H. Tsou, "Perception of forbidden colors in retinally stabilized equiluminant images: an indication of softwired cortical color opponency?" *J. Opt. Soc. Am. A*, vol. 18, no. 10, pp. 2398–2403, Oct 2001.

[19] P.-J. Hsieh and P. U. Tse, "Illusory color mixing upon perceptual fading and filling-in does not result in 'forbidden colors'," *Vision Research*, vol. 46, no. 14, pp. 2251–2258, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0042698905006693

[20] B. Zou, I. S. Utochkin, Y. Liu, and J. M. Wolfe, "Binocularity and visual search—revisited," *Attention, Perception, & Psychophysics*, vol. 79, no. 2, pp. 473–483, 2017.

[21] A. Krekhov, S. Cmentowski, A. Waschk, and J. Krüger, "Deadeye visualization revisited: Investigation of preattentiveness and applicability in virtual environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 547–557, 2020.

[22] T. L. Ooi and Z. J. He, "Sensory eye dominance: Relationship between eye and brain," *Eye and Brain*, vol. 12, p. 25, 2020.

[23] E. Yang, R. Blake, and J. E. McDonald, "A new interocular suppression technique for measuring sensory eye dominance," *Investigative Ophthalmology & Visual Science*, vol. 51, no. 1, pp. 588–593, 2010.

[24] E. Shneor and S. Hochstein, "Eye dominance effects in feature search," *Vision research*, vol. 46, no. 25, pp. 4258–4269, 2006.

[25] ——, "Eye dominance effects in conjunction search," *Vision research*, vol. 48, no. 15, pp. 1592–1602, 2008.

[26] E. Bayle, S. Hourlier, S. Lelandais, C.-A. Salasc, L. Leroy, J. Plantier, and P. Neveu, "Interocular conflict from a monocular augmented reality display: Impact of visual characteristics on performance," *PLOS ONE*, vol. 16, no. 9, pp. 1–24, 09 2021.

[27] M. P. Browne, M. Winterbottom, and R. E. Patterson, "Performance and comfort of monocular head-mounted displays in flight simulators," in *Head- and Helmet-Mounted Displays XV: Design and Applications*, P. L. Marasco and P. R. Havig, Eds., vol. 7688, International Society for Optics and Photonics. SPIE, 2010, pp. 99 – 106.

[28] J. Grudin, "Rivalry and interference with a head-mounted display," *ACM Trans. Comput.-Hum. Interact.*, vol. 9, no. 3, p. 238–251, sep 2002.

[29] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision*, vol. 8, no. 3, pp. 33–33, 03 2008.

[30] R. Xiao and H. Benko, "Augmenting the field-of-view of head-mounted displays with sparse peripheral displays," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1221–1232.

[31] M. A. Livingston, J. L. Gabbard, J. E. Swan, C. M. Sibley, and J. H. Barrow, *Basic Perception in Head-Worn Augmented Reality Displays*. New York, NY: Springer New York, 2013, pp. 35–65.

[32] J. L. Gabbard, J. E. Swan, J. Zedlitz, and W. W. Winchester, "More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces," in *2010 IEEE Virtual Reality Conference (VR)*, 2010, pp. 79–86.

[33] T. Langlotz, M. Cook, and H. Regenbrecht, "Real-time radiometric compensation for optical see-through head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 11, pp. 2385–2394, 2016.

[34] M. Speicher, B. D. Hall, and M. Nebeling, *What is Mixed Reality?* New York, NY, USA: Association for Computing Machinery, 2019, p. 1–15. [Online]. Available: https://doi.org/10.1145/3290605.3300767

[35] R. T. Azuma, "A survey of augmented reality," *Presence: teleoperators & virtual environments*, vol. 6, no. 4, pp. 355–385, 1997.

[36] Y. Itoh, T. Langlotz, J. Sutton, and A. Plopski, "Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays," *ACM Comput. Surv.*, vol. 54, no. 6, jul 2021.

[37] D. Van Krevelen and R. Poelman, "A survey of augmented reality technologies, applications and limitations," *International journal of virtual reality*, vol. 9, no. 2, pp. 1–20, 2010.

[38] J. L. Gabbard, J. E. Swan, and D. Hix, "The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality," *Presence*, vol. 15, no. 1, pp. 16–32, 2006.

[39] N. Sugano, H. Kato, and K. Tachibana, "The effects of shadow representation of virtual objects in augmented reality," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, 2003, pp. 76–83.

[40] W. Lee and W. Woo, "Real-time color correction for marker-based augmented reality applications," in *International Workshop on Ubiquitous Virtual Reality*, 2009, pp. 32–25.

[41] N. Hassani and M. J. Murdoch, "Investigating color appearance in optical see-through augmented reality," *Color Research & Application*, vol. 44, no. 4, pp. 492–507, 2019.

[42] M. A. Livingston, J. H. Barrow, and C. M. Sibley, "Quantification of contrast sensitivity and color perception using head-worn augmented reality displays," in *2009 IEEE Virtual Reality Conference*, 2009, pp. 115–122.

[43] J. L. Gabbard, J. E. Swan, and A. Zarger, "Color blending in outdoor optical see-through ar: The effect of real-world backgrounds on user interface color," in *2013 IEEE Virtual Reality (VR)*, 2013, pp. 157–158.

[44] S. Rothe, D. Buschek, and H. Hußmann, "Guidance in cinematic virtual reality-taxonomy, research status and challenges," *Multimodal Technologies and Interaction*, vol. 3, no. 1, p. 19, 2019.

[45] L. T. Nielsen, M. B. Møller, S. D. Hartmeyer, T. C. M. Ljung, N. C. Nilsson, R. Nordahl, and S. Serafin, "Missing the point: An exploration of how to guide users' attention during cinematic virtual reality," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, ser. VRST '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 229–232.

[46] D. Lange, T. C. Stratmann, U. Gruenefeld, and S. Boll, "Hivefive: Immersion preserving attention guidance in virtual reality," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–13.

[47] J. Orlosky, P. Kim, K. Kiyokawa, T. Mashita, P. Ratsamee, Y. Uranishi, and H. Takemura, "Vismerge: Light adaptive vision augmentation via spectral and temporal fusion of non-visible light," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2017, pp. 22–31.

[48] H. Zhang, X. Cao, and S. Zhao, "Beyond stereo: An exploration of unconventional binocular presentation for novel visual experience," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 2523–2526. [Online]. Available: https://doi.org/10.1145/2207676.2208638

[49] J. Lee, T. Zhan, and S.-T. Wu, "Prospects and challenges in augmented reality displays," *Virtual Reality & Intelligent Hardware*, vol. 1, p. 77, Feb. 2019.

[50] N. K. Logothetis, D. A. Leopold, and D. L. Sheinberg, "What is rivalling during binocular rivalry?" *Nature*, vol. 380, no. 6575, pp. 621–624, Apr. 1996.

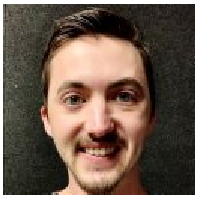
[51] C. Casco, A. Grieco, E. Giora, and M. Martinelli, "Saliency from orthogonal velocity component in texture segregation," *Vision Research*, vol. 46, no. 6, pp. 1091–1098, 2006.

[52] R. T. Pramod and S. P. Arun, "Features in visual search combine linearly," *Journal of Vision*, vol. 14, no. 4, pp. 6–6, 04 2014.

[53] M. Turatto and G. Galfano, "Color, form and luminance capture attention in visual search," *Vision Research*, vol. 40, no. 13, pp. 1639–1643, 2000.

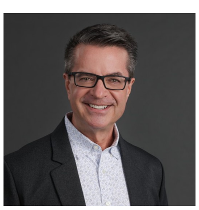
[54] S. Engmann, B. M. 't Hart, T. Sieren, S. Onat, P. König, and W. Einhäuser, "Saliency on a natural scene background: Effects of color and luminance contrast add linearly," *Attention, Perception, & Psychophysics*, vol. 71, no. 6, pp. 1337–1352, Aug. 2009.

[55] S. Etchebehere and E. Fedorovskaya, "On the role of color in visual saliency," *Electronic Imaging*, vol. 2017, no. 14, pp. 58–63, 2017.



Austin Erickson (Graduate Student Member, IEEE) is a doctoral candidate studying computer science at the University of Central Florida. He received his Bachelor of Science degree in computer science from the University of South Florida in July of 2018, and received his Master of Science degree in computer science from the University of Central Florida in May of 2021. His research interests include augmented/virtual reality, visual perception, human computer interaction, and 3D user interfaces.

Gerd Bruder (Member, IEEE) is a Research Assistant Professor for virtual and augmented reality at the Institute for Simulation and Training at the University of Central Florida. He held prior positions as a postdoctoral scholar at the University of Hamburg (2014–2016) and University of Würzburg (2011–2014). He received his Habilitation in Computer Science from the University of Hamburg in 2017 and his Ph.D. in Computer Science from the University of Münster in 2011. His research interests include perception and cognition, virtual humans, social presence, locomotion, and 3D user interfaces.



Gregory F. Welch (Fellow, IEEE) Prof. Welch is a computer scientist and engineer, with appointments in Nursing, Computer Science (CS), and the Institute for Simulation & Training at the University of Central Florida. Welch earned his B.S. in Electrical Engineering Technology from Purdue University (*Highest Distinction*), and his M.S. and Ph.D. in CS from UNC Chapel Hill. Previously, he was a research professor at UNC, and worked for NASA and Northrop. His research interests include human-computer interaction, virtual and augmented reality, and medical applications.