Effects of Dark Mode on Visual Fatigue and Acuity in Optical See-Through Head-Mounted Displays

Kangsoo Kim
University of Central Florida
Orlando, Florida
kangsoo.kim@ucf.edu

Austin Erickson
University of Central Florida
Orlando, Florida
ericksona@knights.ucf.edu

Alexis Lambert
University of Central Florida
Orlando, Florida
lamberta2@knights.ucf.edu

Gerd Bruder
University of Central Florida
Orlando, Florida
bruder@ucf.edu

Gregory F. Welch
University of Central Florida
Orlando, Florida
welch@ucf.edu

ABSTRACT

Light-on-dark color schemes, so-called "Dark Mode," are becoming more and more popular over a wide range of display technologies and application fields. Many people who have to look at computer screens for hours at a time, such as computer programmers and computer graphics artists, indicate a preference for switching colors on a computer screen from dark text on a light background to light text on a dark background due to perceived advantages related to visual comfort and acuity, specifically when working in low-light environments.

In this paper, we investigate the effects of dark mode color schemes in the field of optical see-through head-mounted displays (OST-HMDs), where the characteristic "additive" light model implies that bright graphics are *visible* but dark graphics are *transparent*. We describe a human-subject study in which we evaluated a normal and inverted color mode in front of different physical backgrounds and among different lighting conditions. Our results show that dark mode graphics on OST-HMDs have significant benefits for visual acuity, fatigue, and usability, while user preferences depend largely on the lighting in the physical environment. We discuss the implications of these effects on user interfaces and applications.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in visualization**; • **Computing methodologies** → **Mixed / augmented reality**.

KEYWORDS

Augmented Reality; Optical See-Through Display; Dark Mode; Eye Fatigue; Visual Acuity; User Experience

1 INTRODUCTION

A large body of literature in the field of human-computer interaction focused on colors and light in user interfaces among different display technologies, and evaluated their benefits and drawbacks for different types of tasks. Dark themed user interfaces and so-called "Dark Modes" are gaining popularity in recent years in that they are characterized by a reversal of the prevalent color choices in most user interfaces, i.e., the light-on-dark color scheme that light (e.g., white) font colors are shown on a dark (e.g., black) background. Normal and inverted color choices were investigated over a wide range of display technologies and environments, and were linked to effects of legibility, aesthetics, energy savings, semantic effects, and emotions [5, 24, 25].

Effects of different color modes depend on the display technology used [5]. In this scope, augmented reality (AR) displays present an under-explored domain [36]. In particular, optical see-through (OST) displays and their corresponding light models change the way light and dark colors are perceived compared to traditional screens [14].

OST displays employing beam splitters can be classified as *additive*, which means that they have their own light source, or *subtractive*, meaning that they filter light from an external light source. The most popular current OST head-mounted displays (HMDs), such as the Microsoft HoloLens, use an additive design, which is characterized by "color blending" as they mix (a) light reaching the user's eyes from the physical environment as well as (b) light emitted by the display [12]. The practical effect is that these displays can only add light but not "take light away" from the user's view, which means that dark colors can not be induced but depend entirely on a dark background in the physical environment when looking through the display. Technological solutions to these issues are difficult to realize and will not be available to consumers in the foreseeable future [18, 31, 35]. Hence, with AR annotations on OST-HMDs, such as black text on a white background, the white pixels

match the brightness of the display and the black pixels match the color of the physical background behind the pixels [16].

It is our hypothesis that common AR annotations on OST-HMDs, involving light/dark foreground colors on a dark/light background, can have a significant impact on users' visual acuity and fatigue. In this paper, we discuss and investigate these challenges on the example of a Microsoft HoloLens. We asked participants in a human-subject study to read AR annotations with different color modes and to complete visual acuity tests with AR annotations in front of different backgrounds and with different lighting conditions. We assessed participants' visual fatigue, acuity, and preferences.

In particular, we investigated the following research questions:

RQ1 Are there subjective or objective benefits of dark mode color schemes with OST-HMDs with respect to visual fatigue and acuity?

RQ2 Do users subjectively prefer normal or dark mode color schemes with OST-HMDs?

RQ3 Do users' preferences match the objective benefits or drawbacks of the modes?

This paper is structured as follows. Section 2 presents an overview of related work. Section 3 describes the human-subject study. The results are presented in Section 4 and discussed in Section 5. Section 6 concludes the paper and discusses future research.

2 RELATED WORK

Computer displays usually strive to present information with a high signal-to-noise ratio, in particular when presenting text to readers, which emphasizes the benefits of strong luminance differences instead of chromatic differences between the foreground and background. In the early age of electronic display technology when cathode-ray tube (CRT) monitors were prevalent, light-on-dark color scheme interfaces, i.e., light text on a dark background, were common because the text on the monitors was displayed by the electron beam hitting the phosphorous material for luminescence that is normally dark in the normal state. However, as the dark-on-light color scheme, i.e., dark text on a light background, was introduced in WYSIWYG editing systems to simulate ink on paper in the real world, it has been dominant in many computer user interfaces. Presenting dark text on a light background is usually referred to as *positive contrast*, which goes back to the signal processing theory, where the peak-to-peak contrast (or Michelson contrast [28]) measures the ratio between the spread and the sum of two luminances. This ratio is defined as $c = \frac{L_b - L_t}{L_b + L_t}$ with text luminance L_t and background luminance L_b, which is negative if $L_b < L_t$. While both positive and negative contrast conditions can provide the same theoretical peak-to-peak contrast ratio, a large body of literature focused on identifying benefits of one of them over the other for different display technologies and use cases.

Multiple studies have found that *positive contrast* has benefits when the goal is to read text on computer screens [1, 9, 34]. More recent studies investigated the causes of these benefits. Taptagaporn and Saito observed that participants developed a smaller *pupil diameter* when they used a positive contrast display compared to a negative contrast display [32]. A small pupil diameter is known to increase the quality of the retinal image with greater depth of field and less spherical aberration, and it is largely affected by the

amount of light reaching the observer's eyes. Buchner et al. investigated the display luminance in positive and negative contrast modes, showing that it is usually *higher* in positive contrast modes, e.g., when dark text is presented on a light background [5], which can be traced back to the ratio of screen space filled by (dark) letters or the (light) background. They further performed a study showing that, indeed, the amount of luminance had a dominant effect on performance while reading, but there was no difference between positive and negative contrast modes if the overall luminance was equivalent. In other words, by increasing the lightness of letters on a dark background they created the same effect that dark letters had on a moderately light background.

While the benefits of the positive contrast mode originate in the increased display luminance, this is not always desirable. For instance, the automotive industry has a long history of designing in-car displays and illumination for daytime and nighttime use. While positive contrast could be beneficial in terms of reading text on in-car displays independently of the environment lighting conditions, increasing the amount of light reaching the driver's eyes can have negative effects during the night, since it reduces the dark adaptation of the driver's eyes and thus their ability to perceive obstacles or people in low light road conditions [26]. Modern in-car displays thus usually switch to a "night mode" when it gets dark outside, which is characterized by a switch from a positive contrast (daytime) to a negative contrast (nighttime) mode. Also, human circadian physiology and cognitive performance can be influenced by different displays [7]. Higuchi et al. found that performing a task with a bright display influences the nocturnal melatonin concentration and other physiological indicators of the human biological clock [15].

In modern life, people spend an increasing amount of time in front of computer screens, and experience various ocular symptoms, such as eyestrain, tired eyes, and sensitivity to bright lights and eye discomfort, which are referred to as computer vision syndrome (CVS) [3]. Various recommendations have been made with regard to luminance values for background and characters. Campbell and Durden emphasized that individual users should be able to adjust the brightness of the computer devices to adjust the luminance and contrast depending upon the time and the ambient lighting of the workplace [8]. In this sense, researchers and practitioners have developed automatic color adjustment programs or dark mode color schemes to protect the user's eyes and increase the comfort level, e.g., f.lux[1]. They further allow for these settings to be fine-tuned by users to their specific preferences. Due to the popularity of such features among users, many companies adopted the dark mode interface design scheme in their hardware and software. For example, Apple included the feature of a dark mode setting that could be applied to adjust the coloration and brightness of all core applications on the device to a darker format with the release of their operating system Mojave[2].

There have been many studies about the effects of different displays on visual fatigue and acuity, e.g., 3D displays and virtual reality (VR) headsets [20, 21, 23, 37]. Even in the domain of AR

---

[1] f.lux, a cross-platform software that adjusts a display's color temperature accordingly (https://justgetflux.com/)

[2] Apple, "How to Use Dark Mode on your Mac" (https://support.apple.com/en-us/HT208976).

research, researchers investigated the effects of real background patterns and focal distance on visual fatigue and acuity [13, 27]. However, we are not aware of any work on positive or negative contrast modes in the field of AR, specifically with respect to text-based AR annotations. We see parallels between in-car heads-up displays and current-state OST-HMDs in the use of additive display designs (see Section 1), and the overall desire not only to ensure legibility of the displayed text but also to retain natural viewing of the physical environment behind the display without inducing severe visual fatigue.

3 EXPERIMENT

In this section we describe the experiment that we conducted to investigate the three research questions stated in Section 1. Participants were asked to read text in AR on a HoloLens OST-HMD under different vision modes, and we asked them to complete visual acuity tests as well as rate their visual fatigue, subjective experience and preference of these vision modes.

3.1 Participants

We recruited 19 participants for our experiment; ten male and nine female (ages 18 to 41, M=25.47, SD=5.93). The participants were members of the local university community. All of the participants had normal or corrected-to-normal vision; four participants wore glasses during the experiment, and four wore contact lenses. None of the participants reported known visual or vestibular disorders, such as color or night blindness, dyschromatopsia, or a displacement of balance. We ensured the normal condition of the participants' eyes by measuring the Ocular Surface Disease Index (OSDI) [29], which consists of 12 questions evaluating the frequency of dry eye disease symptoms over the preceding week. All 19 participants were categorized as normal with an OSDI score that is less than 12 in the range of 0–100. 18 participants reported that they had used a VR or AR HMD in the past, and four of them rated themselves as frequent users, having used HMDs on more than ten separate occasions. We asked participants to rate their current preference and usage of dark mode and inverted color schemes on their computers and mobile devices before the experiment. Two participants used these features whenever they were available, seven participants used these modes frequently, nine used these modes occasionally, and one never made use of these modes.

3.2 Materials

3.2.1 AR Stimuli and Vision Modes. For the presentation of the visual stimuli, we used a Microsoft HoloLens so that participants could see the AR visual stimuli, which were displayed in front of them (Figure 1 and Figure 2). As a widely-used OST-HMD, the HoloLens provides an augmented field of view of circa 30 degrees horizontally by 17 degrees vertically in the center of the total human visual field. The resolution is 1268×720 pixels per eye. The HoloLens leverages SLAM-based tracking [6] to localize itself with respect to the physical environment.

For the rendering of the visual stimuli, we used the Unity game engine and its integration with the HoloLens. We chose AR textual annotations registered as planar objects ("holograms") in the laboratory space that consisted of either black text on a white background
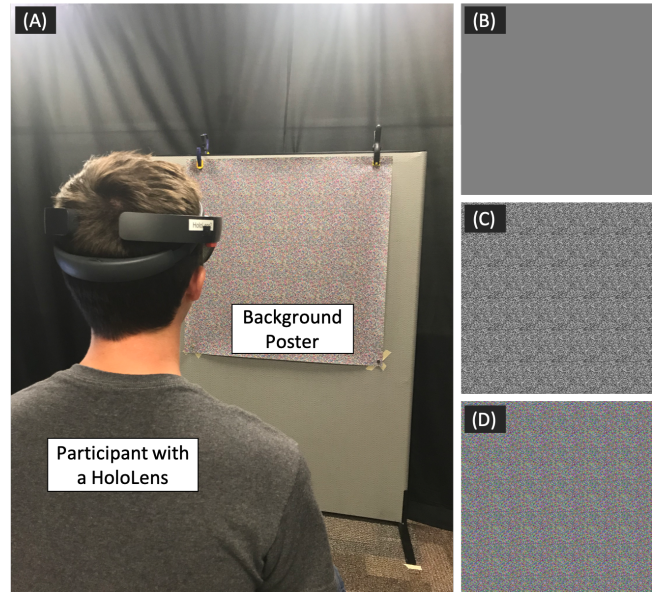

Figure 1: Experimental setup and the images for the background posters. (A) A participant wearing the HoloLens, seated in front of the background poster. (B) The image used for the solid grey background poster. (C) The image used for the lightness distortion background poster. (D) The image used for the chromatic distortion background poster.

or white text on a black background. Participants in the study were positioned 1.52 meters away from the annotations, which were presented at the participant's eye height to avoid inclination conflicts. We prepared four reading passages extracted from Pearson Test of English Read Aloud Practice Questions[3] in a 2×2 grid text board in a size of 72 by 72 centimeters (Figure 2).

We further developed an AR version of a common visual acuity test chart similar to a Golovin–Sivtsev Table with Landolt C characters [10, 11, 33], which are characterized by circles with a missing piece on either of four sides (Figure 3). We implemented a randomized version of this test, where each trial resulted in different orientations of these circles. The chart had a physical (registered) size of 36 by 36 centimeters.

Again, the chart either consisted of black circles on a white background or white circles on a black background. The chart was placed at the same distance as the AR annotations.

The two considered vision mode conditions were as follows:

- **Light Mode**: We used a *positive* contrast mode in which the foreground was presented as black and the background as white on the HoloLens.
- **Dark Mode**: We used a *negative* contrast mode in which the foreground was presented as white and the background as black.

---

[3]Pearson Test of English (PTE) Read Aloud Practice Questions (https://pteacademicexam.com/pte-academic-speaking-read-aloud-practice-test-1-sample-exercises/).
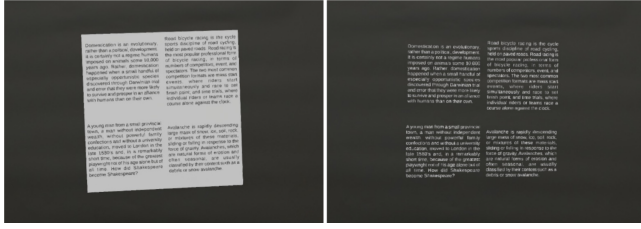
Figure 2: Illustrations showing the AR text-reading task participants had to read during the experiment: the light mode (left) and the dark mode (right).

3.2.2 Physical Environment and Background. We prepared an isolated room, which was surrounded by black curtains, in our laboratory space so that participants were not exposed to other visual stimuli during the study (Figure 1(A)). We created different backgrounds for the experiment by mounting large-scale printed posters on a partition wall in front of the participants (Figure 1(B–D)). The posters were made of 36" × 36" Premium Archival Matte papers. The three considered background conditions were as follows (Figure 1 and Figure 3):

- **Uniform**: Participants perceived a uniform gray background (printed using a pixel intensity of 128 in the range of 0–255).
- **Lightness Distortions**: The background consisted of a mixture of randomly generated gray-scale pixels, impacting the apparent luminance of the stimuli being presented on the OST-HMD.
- **Chromatic Distortions**: The background consisted of a mixture of randomly generated RGB pixels, creating chromatic differences in line with an exaggerated simulation of using an OST-HMD in a cluttered environment.

3.2.3 Physical Lighting. To evaluate the differences between the amount of light in the physical environment, we controlled the overall lighting in the experimental setup. We created a well-lit environment that illuminated the room, and we compared it to a reduced-light environment. The two considered lighting conditions in this experiment were as follows (Figure 3):

- **High Light**: The environment was well-lit due to diffuse indirect ceiling lighting in the room (with 200–270 Lux[4]).
- **Low Light**: The environment was reasonably dark due to dimmed lighting (with 10–12 Lux).

3.3 Methods

We used a full-factorial within-subjects design in this experiment. As described in Section 3.2, the independent variables were as follows:

- **Vision Mode** (*Light Mode, Dark Mode*),
- **Physical Background** (*Uniform, Lightness Distortions, Chromatic Distortions*), and
- **Physical Lighting** (*High Light, Low Light*).

---

[4]Measured by a URCERI Light Meter Digital Illuminance Meter

(A) Uniform background

(B) Lightness distortion background
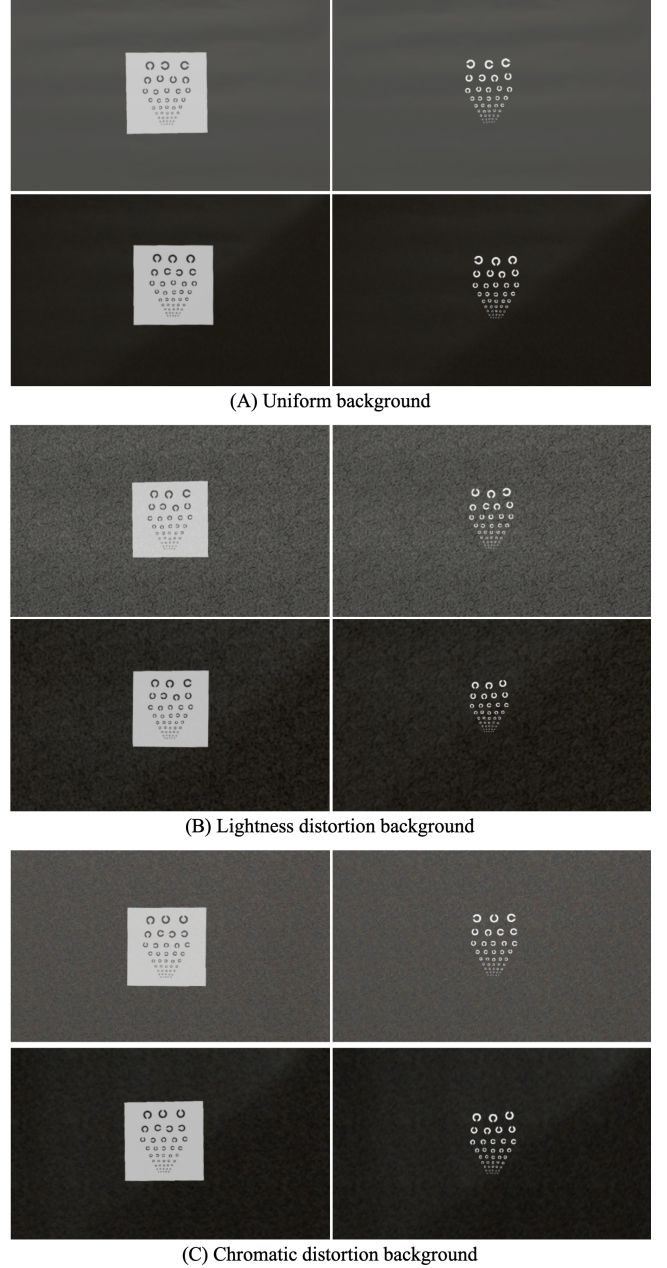

(C) Chromatic distortion background

Figure 3: Illustrations of the different experimental conditions. The left column shows the light mode AR stimuli and the right column is the dark mode stimuli. From top to bottom, the rows show the different backgrounds: (A) uniform, (B) lightness distortion, (C) chromatic distortion. Each in the high-light and low-light physical lighting conditions. These picture were taken from the HoloLens via the device portal.

Table 1 summarizes the conditions. Each participant completed all twelve conditions while we tried to balance the orders with respect to the variables.

Table 1: Study conditions: Each column refers to factors we controlled. The rows indicate the tested combinations.

Vision Mode	Lighting	Background
Light Mode	High	Uniform
		Lightness Distortion
		Chromatic Distortion
	Low	Uniform
		Lightness Distortion
		Chromatic Distortion
Dark Mode	High	Uniform
		Lightness Distortion
		Chromatic Distortion
	Low	Uniform
		Lightness Distortion
		Chromatic Distortion

3.3.1 Procedure. Prior to the experiment trials, participants first were asked to give their informed consent. Afterwards, they received task instructions and the experimenters made sure that they understood the task. Participants performed the interpupillary distance (IPD) calibration on the HoloLens before the experiment, so that the virtual content was rendered correctly in their view. Participants further completed a demographics questionnaire and the OSDI, which evaluated their current level of visual fatigue (see Section 3.1). Participants then started the experimental trials.

At the beginning of each trial, participants were instructed to sit on the designated chair positioned directly in front of the wall which supported the posters (Figure 1). Participants were asked to verify that the positioning of the chart was correct before observing a set of four paragraphs that would be displayed for one minute (Figure 2). During this time, participants were asked to read the paragraphs (which were the same for each trial), and observe how easy or difficult the text was to read while sensing their general preference and fatigue level. After one minute had passed, the participant performed the (randomly generated) visual acuity test, where their accuracy and response time was recorded (Figure 3). Following each trial, participants were asked to complete two questionnaires: a short usability questionnaire and a visual fatigue questionnaire.

After completing four trials with the same physical background poster, participants were asked to further compare the light mode and dark mode AR annotations and choose their preferred option for the displayed lighting and background combination. They were also asked to choose which option they found to be most comfortable, which option they found to be easiest to read, and which option they thought that they performed better on. After answering these questions for both lighting conditions, the background poster was changed for the next set of four trials.

After completing all trials, participants had a brief interview with the experimenter asking their overall perception or feeling about all the tested conditions. Finally they received monetary compensation ($15) and finished the study.

3.3.2 Measures. We collected both objective and subjective measures to understand the benefits or drawbacks of the vision modes under the different background and lighting conditions in AR.

We considered the following dependent variables:

- **Visual Acuity**: As explained in Section 3.2.1, we used a visual acuity test based on a Golovin–Sivtsev Table with Landolt C characters [10, 11, 33]. The acuity is computed

by the number of mistakes that a participant makes when reading from the chart.

- **Visual Fatigue**: To assess increases in visual fatigue that were induced during the trials in the study, we modified the revised convergence insufficiency symptom survey (CISS) [4]. The convergence insufficiency is a common binocular vision disorder, and it is normally associated with symptoms such as visual fatigue and headaches. While their survey asks the frequency of the symptoms in a scale of five degrees (Never–Always), we converted the questions to a 7-point Likert scale (Strongly Disagree–Strongly Agree) considering the task for the experiment.
- **Usability**: We asked participants to rate the usability of the AR annotations after each condition using the short user experience questionnaire (UEQ-S) [30]. While the original UEQ is a semantic differential with 26 items, the UEQ-S consists of only eight items. The UEQ-S focuses on the measurement of the two meta-dimensions, *pragmatic* and *hedonic* qualities, and evaluates the overall usability based on those two quality aspects.
- **Preferences**: We asked participants to indicate their subjective preferences and rank the two vision modes for each of the background and lighting conditions.

We further debriefed the participants and asked them to verbalize additional observations and impressions.

3.4 Hypotheses

Based on the related work, and our study design, we formulated the following hypotheses for the objective and subjective results:

H1 Participants will show higher visual acuity with the light mode AR annotations than using the dark mode.

H2 Participants will show higher visual fatigue with the light mode AR annotations than using the dark mode.

H3 Participants will indicate higher subjective ratings of usability and preference for the dark mode AR annotations in dark physical environments.

4 RESULTS

We used parametric statistical tests to analyze the responses in line with the ongoing discussion in the field of psychology indicating that parametric statistics can be a valid and more informative method for the analysis of combined experimental questionnaire scales with individual ordinal data points measured by questionnaires or coded behaviors [19, 22]. We analyzed the responses with repeated-measures ANOVAs and Tukey multiple comparisons with Bonferroni correction at the 5% significance level. We confirmed the normality with Shapiro-Wilk tests at the 5% level and QQ plots. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly's test indicated that the assumption of sphericity had been violated. We only report the significant effects.

4.1 Visual Acuity

The results for the visual acuity test are shown in Figure 4.

We found a significant main effect for **vision mode** between the light mode ($M = 7.62$, $SD = 1.31$) and the dark mode ($M = 8.32$,
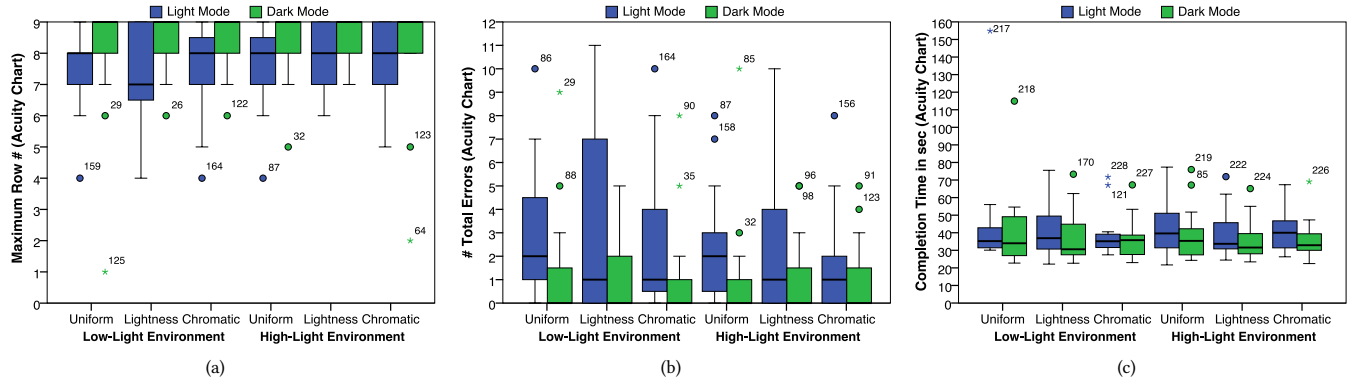
Figure 4: Results for the *visual acuity* tests: (a) maximum row on the acuity chart that could be completed without errors (between 0 and 9; higher is better), (b) total number of errors on acuity chart (lower is better), and (c) completion time for the acuity chart (lower is better).
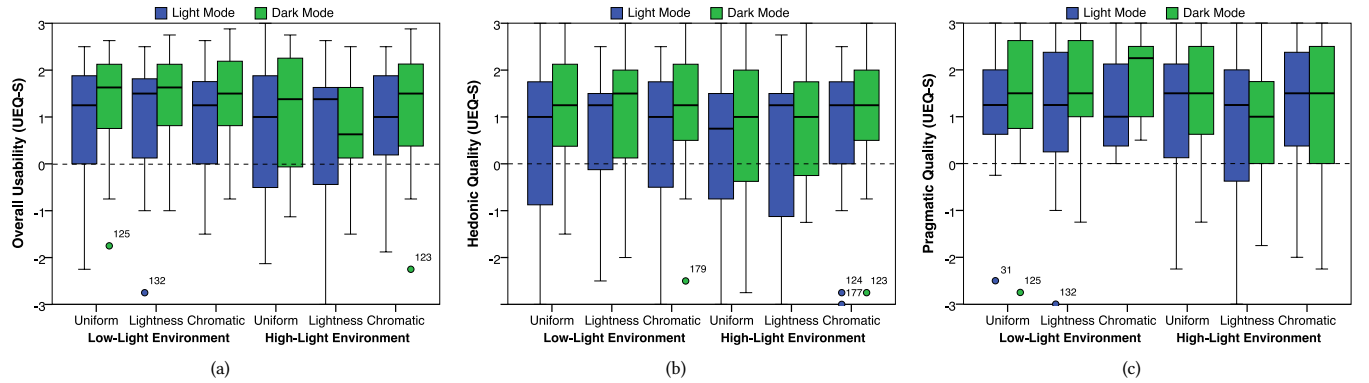

Figure 5: Results for the *usability* estimates using the UEQ-S questionnaire (between -3 and 3; higher is better): (a) overall usability score, (b) hedonic quality, and (c) pragmatic quality.

SD = 1.27) on the maximum row on the visual acuity chart that could be completed without errors, $F(1, 18) = 9.20$, $\boldsymbol{p = 0.007}$, $\eta_p^2 = 0.338$, indicating that participants had a significantly higher visual acuity for the dark mode than the light mode.

We found a significant main effect for **vision mode** between the light mode ($M = 2.52$, $SD = 2.95$) and the dark mode ($M = 1.10$, $SD = 1.94$) on the numbers of errors made in the visual acuity tests, $F(1, 18) = 12.65$, $\boldsymbol{p = 0.002}$, $\eta_p^2 = 0.413$, indicating that participants completed the tests with significantly fewer errors for the dark mode than the light mode.

We found a significant main effect for **physical lighting** between the low-light ($M = 2.04$, $SD = 2.84$) and high-light ($M = 1.57$, $SD = 2.30$) environment on the numbers of errors made in the visual acuity tests, $F(1, 18) = 11.68$, $\boldsymbol{p = 0.003}$, $\eta_p^2 = 0.394$, indicating that participants completed the tests with significantly fewer errors in the high-light environment than in the low-light environment.

We also found a significant main effect for **vision mode** between the light mode ($M = 40.89$ sec, $SD = 16.64$ sec) and the dark mode ($M = 37.18$ sec, $SD = 14.20$ sec) on the completion time of the visual

acuity tests, $F(1, 17) = 16.04$, $\boldsymbol{p = 0.001}$, $\eta_p^2 = 0.485$, indicating that participants completed the tests significantly faster for the dark mode than the light mode.

4.2 Visual Fatigue

The results for visual fatigue (CISS) are shown in Figure 6.

We found a significant main effect for **vision mode** between the light mode ($M = 21.07$, $SD = 19.39$) and dark mode ($M = 16.69$, $SD = 16.95$) on CISS visual fatigue scores, $F(1, 18) = 15.22$, $\boldsymbol{p = 0.001}$, $\eta_p^2 = 0.458$, indicating that participants' visual fatigue was significantly lower with the dark mode than with the light mode.

4.3 Usability

The results for usability (UEQ-S) are shown in Figure 5.

We found a significant main effect for **vision mode** on *overall usability*, $F(1, 17) = 9.17$, $\boldsymbol{p = 0.008}$, $\eta_p^2 = 0.350$, specifically on *hedonic quality*, $F(1, 17) = 10.22$, $\boldsymbol{p = 0.005}$, $\eta_p^2 = 0.375$, and a trend for *pragmatic quality*, $F(1, 17) = 4.17$, $p = 0.057$, $\eta_p^2 = 0.197$. The
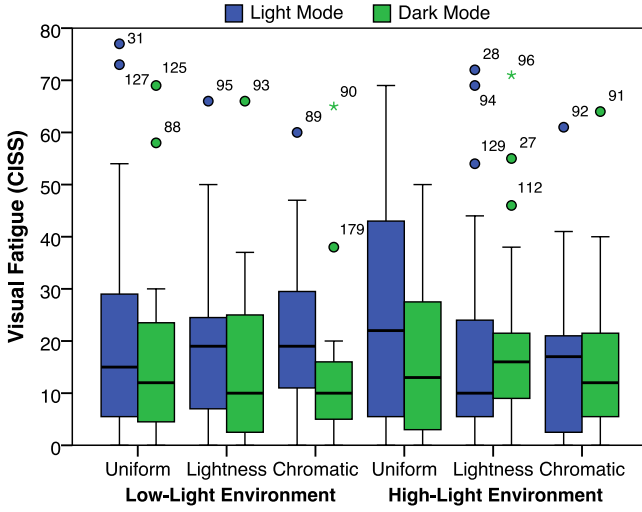
Figure 6: Results for the *visual fatigue* estimates using the CISS questionnaire (lower is better).

results indicate that participants rated usability as significantly higher for the dark mode than the light mode.

We found a significant main effect for **physical lighting** on *overall usability*, $F(1, 17) = 7.00$, ***p = 0.017***, $\eta_p^2 = 0.292$, specifically on *pragmatic quality*, $F(1, 17) = 6.25$, ***p = 0.023***, $\eta_p^2 = 0.269$, and a trend for *hedonic quality*, $F(1, 17) = 3.72$, $p = 0.071$, $\eta_p^2 = 0.180$. The results indicate that participants rated usability as significantly higher for the low-light physical environment than the high-light environment.

We found a significant interaction effect between **physical lighting** and **vision mode** on *overall usability*, $F(1, 17) = 6.85$, ***p = 0.018***, $\eta_p^2 = 0.287$, specifically on *pragmatic quality*, $F(1, 17) = 4.89$, ***p = 0.041***, $\eta_p^2 = 0.223$, and on *hedonic quality*, $F(1, 17) = 6.97$, ***p = 0.017***, $\eta_p^2 = 0.291$. Multiple comparisons (all $p < 0.05$) showed that *overall usability*, *pragmatic quality*, and *hedonic quality* were all significantly higher for the dark mode than the light mode in the low-light environment. We further found that the three measures were significantly higher for the low-light environment than the high-light environment for the dark mode. Last but not least, we found that *overall usability* and *hedonic quality* (only a trend with $p = 0.082$ for *pragmatic quality*) were significantly higher for the dark mode in the low-light environment than the light mode in the high-light environment.

4.4 Subjective Preferences

The subjective preferences of our participants are shown in Figure 7.

We found significant main effects for **physical lighting** on *overall preference*, $F(1, 18) = 6.09$, ***p = 0.024***, $\eta_p^2 = 0.253$, on *visual comfort*, $F(1, 18) = 7.44$, ***p = 0.014***, $\eta_p^2 = 0.292$, on *easy to read*, $F(1, 18) = 9.21$, ***p = 0.007***, $\eta_p^2 = 0.338$, and on *perceived performance*, $F(1, 18) = 19.36$, ***p < 0.001***, $\eta_p^2 = 0.518$. The results show that the dark mode is mainly the preferred choice in the low-light environment and less so in the high-light environment.

We also found a significant main effect for **background** on *overall preference*, $F(1.51, 27.17) = 5.78$, ***p = 0.013***, $\eta_p^2 = 0.243$. Post-hoc tests (all $p < 0.05$) showed a significantly higher preference for the dark mode for the uniform background than the lightness/chromatic backgrounds for *overall preference*. They further showed a significantly higher preference for the dark mode for the uniform background than the lightness background for *visual comfort* and for *easy to read*.

5 DISCUSSION

In this section, we summarize the main findings and discuss implications for the use of AR user interfaces with current-state OST-HMDs.

5.1 Dark Mode Improves Visual Acuity

In contrast to our Hypothesis **H1**, we found that participants had a significantly higher visual acuity for the dark mode than the light mode. They were able to complete significantly more rows on the visual acuity test chart without errors for the dark mode. Moreover, they also made significantly fewer overall errors on that test for the dark mode conditions, and it also took them significantly less time to complete the test.

This result is interesting as it implies that visual details such as text are easier to see on OST-HMDs if they are presented in light colors over the background (i.e., dark mode) instead of the traditional approach on computer screens, where the details are dark and the space around them is illuminated (i.e., light mode). Since the light mode approach illuminates a screen area on an OST-HMD with gaps of non-illuminated pixels (dark transparent pixels) for the details on the OST-HMD, we assumed that these might be prone to an influence of complex physical backgrounds shining through those gaps.

We also found that participants made significantly fewer errors on the visual acuity test in the high-light physical environment compared to the low-light environment. However, this is in line with the literature and not surprising as it can be explained by the well-known relationship between environment luminance and pupil diameter (see Section 2).

5.2 Dark Mode Reduces Visual Fatigue

In line with our Hypothesis **H2**, we found for the CISS visual fatigue results that participants' visual fatigue was significantly lower for the dark mode than the light mode. This result is associated with the findings of prior research that showed the positive effects of the dark mode on visual fatigue levels. While the light mode for the present study can simulate real texts printed on a physical paper, e.g., dark ink letters on a white paper, it makes the display emit high luminance, which can cause fatigue [3, 15]. For example, Benedetto et al. [2] conducted a study to investigate visual fatigue in terms of two different levels of screen luminance and two levels of ambient lighting, and showed that reading under high levels of screen luminance increases visual fatigue.

(a) Overall Preference   (b) Visual Comfort   (c) Easy to Read   (d) Perceived Performance
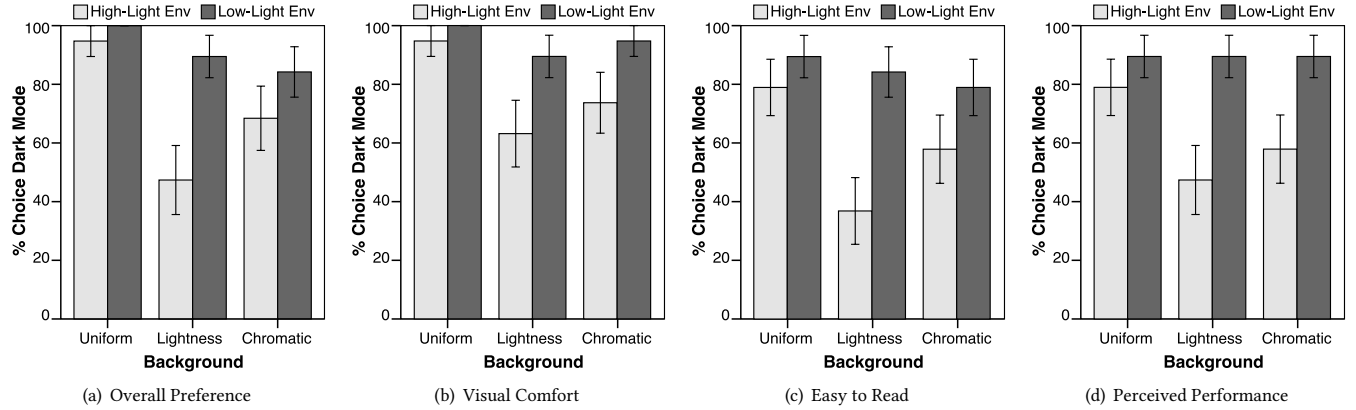
Figure 7: Subjective preferences: Percentages indicating how often participants chose the dark mode over the light mode for the different *physical lighting* and *background* conditions. The plots indicate (a) which of them were overall preferred, (b) perceived as more visually comfortable, (c) perceived as easier to read, and (d) perceived as resulting in higher performance (fast and accurate reading).

5.3 Dark Mode Improves Usability and Preference in Dark Environments

In line with our Hypothesis **H3**, we found that participants indicated a clear overall preference for the dark mode over the light mode. Participants' responses suggest that this goes back to subjective impressions of higher visual comfort with the dark mode, an overall sense of making it easier to read, and the impression that the dark mode increased their performance in AR during the experiment. They further indicated that the dark mode significantly increased the overall usability of AR annotations as well as their hedonic and pragmatic qualities.

We found that participants preferred the dark mode mainly in the low-light environment and less so in the high-light environment. In particular, as shown in Figure 7, participants indicated a balanced preference for either light or dark modes in the high-light environment with a more complex background (with chromatic or lightness distortions), which, arguably, might be more ecologically valid than a uniform background. We would also like to point out that these preferences might further shift towards the light mode in situations with even more environmental light. Both physical lighting setups in our study were designed for typical room interior lighting found in office environments.

5.4 Limitations

As discussed in Section 1, our results are specific to the additive light model used in current-state OST-HMDs. Due to the peculiarities of OST displays, we are making no claims that these results would transfer to video see-through or immersive VR displays, or that these findings would still hold true if technological solutions to the color blending of OST displays become feasible [18, 31, 35].

However, it is likely that strong subjective preferences or a user's sense of gained benefits of the light or dark mode might transcend different display technologies. As listed in Section 3.1, our participants were roughly split into one half who used dark mode graphics occasionally in their daily life and the other half who used them

frequently. We observed no effects of these general tendencies on the results in this study.

6 CONCLUSION AND FUTURE WORK

In this paper, we presented an analysis of positive and negative contrast modes, also known as light and dark modes, with OST-HMDs based on additive light models. We presented a human-subject study that we conducted to understand objective benefits and participants' subjective preferences of these vision modes under different physical lighting and background conditions. We showed that dark mode color schemes used for AR annotations significantly increased visual acuity and reduced visual fatigue, and was overall preferred by participants, especially in low-light physical environments and with complex backgrounds.

In future work, we believe that it is important to investigate related chromatic differences with OST-HMDs, and to investigate vision modes also with video see-through and immersive VR displays. VR/AR displays are well-suited for the simulation of different vision modes and lighting conditions. As visualization and evaluation have become major research areas in AR [17], we believe that our results could lead to a better understanding of human augmented vision across disciplines and technologies ranging from night vision goggles to ambient light displays.

REFERENCES

[1] D. Bauer and C. R. Cavonius. 1990. *Ergonomic Aspects of Visual Display Terminals*. London: Taylor & Francis, Chapter Improving the legibility of visual display units through contrast reversal.

[2] S. Benedetto, A. Carbone, V. Drai-Zerbib, M. Pedrotti, and T. Baccino. 2014. Effects of luminance and illuminance on visual fatigue and arousal during digital reading. *Computers in Human Behavior* 41 (2014), 112–119.

[3] C. Blehm, S. Vishnu, A. Khattak, S. Mitra, and R. W. Yee. 2005. Computer Vision Syndrome: A Review. *Survey of Ophthalmology* 50, 3 (2005), 253–262.

[4] E. J. Borsting, M. W. Rouse, G. L. Mitchell, M. Scheiman, S. A. Cotter, J. Cooper, M. T. Kulp, R. London, and CITT Group. 2003. Validity and Reliability of the Revised Children Aged 9 to 18 Years. *Optometry & Vision Science* 80, 12 (2003), 832–838.

[5] A. Buchner, S. Mayr, and M. Brandt. 2009. The advantage of positive text-background polarity is due to high display luminance. *Ergonomics* 52, 7 (2009), 882–886.

[6] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. Leonard. 2016. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics* 32, 6 (2016), 1309–1332.

[7] C. Cajochen, S. Frey, D. Anders, J. Späti, M. Bues, A. Pross, R. Mager, A. Wirz-Justice, and O. Stefani. 2011. Evening exposure to a light-emitting diodes (LED)-backlit computer screen affects circadian physiology and cognitive performance. *Journal of Applied Physiology* 110, 5 (2011), 1432–1438.

[8] F. W. Campbell and K. Durden. 1983. The Visual Display Terminal Issue: A Consideration of Its Physiological, Psychological and Clinical Background. *Ophthalmic and Physiological Optics* 3, 2 (1983), 175–192.

[9] A. H. S. Chan and P. S. K. Lee. 2005. Effect of display factors on Chinese reading times, comprehension scores and preferences. *Behaviour and Information Technology* 24 (2005), 81–91.

[10] International Organization for Standardization (ISO). 1994. ISO 5725-2:1994 - Accuracy (Trueness and Precision) of Measurement Methods and Results. Basic Methods for the Determination of Repeatability and Reproducibility of a Standard Measurement Method.

[11] International Organization for Standardization (ISO). 2009. 8597:1994 and 8596:2009 - Ophthalmic Optics. Visual Acuity Testing. Standard Optotype and its Presentation.

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

[13] J. L. Gabbard, D. G. Mehra, and J. E. Swan. 2019. Effects of ar display context switching and focal distance switching on human performance. *IEEE Transactions on Visualization and Computer Graphics* 25, 6 (2019), 2228–2241.

[14] J. Harris. 2014. *Sensation and Perception*. SAGE.

[15] S. Higuchi, Y. Motohashi, Y. Liu, M. Ahara, and Y. Kaneko. 2015. Effects of VDT tasks with a bright display at night on melatonin, core temperature, heart rate, and sleepiness. *Journal of Applied Physiology* 94, 5 (2015), 1773–1776.

[16] W. Huang, L. Alem, and M. A. Livingston. 2012. *Human Factors in Augmented Reality Environments*. Springer Science & Business Media.

[17] K. Kim, M. Billinghurst, G. Bruder, H. Duh, and G. F. Welch. 2018. Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017). *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 24, 11 (2018), 2947–2962.

[18] K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods. 2003. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *Proceeding of the International Symposium on Mixed and Augmented Reality (ISMAR)*. 133–141.

[19] T. R. Knapp. 1990. Treating ordinal scales as interval scales: an attempt to resolve the controversy. *Nursing research* 39, 2 (1990), 121–123.

[20] F. L. Kooi and A. Toet. 2004. Visual comfort of binocular and 3D displays. *Displays* 25, 2–3 (2004), 99–108.

[21] K. Krösl, D. Bauer, M. Schwärzler, H. Fuchs, G. Suter, and M. Wimmer. 2018. A VR-based user study on the effects of vision impairments on recognition distances of escape-route signs in buildings. *Visual Computer* 34, 6-8 (2018), 911–923.

[22] W. M. Kuzon Jr, M. G. Urbanchek, and S. McCabe. 1996. The seven deadly sins of statistical analysis. *Annals of plastic surgery* 37, 3 (1996), 265–272.

[23] M. Lambooij, M. Fortuin, W. Ijsselsteijn, B. Evans, and I. Heynderickx. 2010. Measuring visual fatigue and visual discomfort associated with 3-D displays. *Journal of the Society for Information Display* 18, 11 (2010), 931–943.

[24] D. Löffler, L. Giron, and J. Hurtienne. 2017. Night Mode, Dark Thoughts: Background Color Influences the Perceived Sentiment of Chat Messages. In *Proceeding of INTERACT*. 184–201.

[25] L. W. MacDonald. 1999. Using color effectively in computer graphics. *IEEE Computer Graphics and Applications* 19, 4 (1999), 20–35.

[26] S. Mayr and A. Buchner. 2010. After-effects of TFT-LCD display polarity and display colour on the detection of low-contrast objects. *Ergonomics* 53, 7 (2010), 914–925.

[27] C. Merenda, M. Smith, J. Gabbard, G. Burnett, and D. Large. 2016. Effects of real-world backgrounds on user interface color naming and matching in automotive AR HUDs. In *IEEE VR 2016 Workshop on Perceptual and Cognitive Issues in AR (PERCAR)*. 1–6.

[28] A. A. Michelson. 1927. *Studies in optics*. University Press, Chicago.

[29] R. M. Schiffman, M. D. Christianson, G. Jacobsen, J. D. Hirsch, and B. L. Reis. 2000. Reliability and validity of the ocular surface disease index. *Archives of Ophthalmology* 118, 5 (2000), 615–621.

[30] M. Schrepp, A. Hinderks, and J. Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 6 (2017), 103.

[31] S. K. Sridharan, J. D. Hincapie-Ramos, D. R. Flatla, and P. Irani. 2013. Color Correction for Optical See-Through Displays Using Display Color Profiles. In *Proceeding of the ACM Symposium on Virtual Reality Software and Technology (VRST)*. 231–240.

[32] S. Taptagaporn and S. Saito. 1990. How display polarity and lighting conditions affect the pupil size of VDT operators. *Ergonomics* 33 (1990), 201–208.

[33] H. K. Walker, W. D. Hall, and J. W. Hurst. 1990. *Clinical Methods: The History, Physical, and Laboratory Examinations*. Butterworths, Boston.

[34] A.-H. Wang, J.-J. Fang, and C.-H. Chen. 2003. Effects of VDT leading-display design on visual performance of users in handling static and dynamic display information dual-tasks. *International Journal of Industrial Ergonomics* 32 (2003), 93–104.

[35] C. Weiland, A.-K. Braun, and W. Heiden. 2009. Colorimetric and photometric compensation for optical see-through displays. *Universal Access in Human-Computer Interaction* 5615 (2009), 603–612.

[36] G. Welch, G. Bruder, P. Squire, and R. Schubert. 2019. *Anticipating Widespread Augmented Reality: Insights from the 2018 AR Visioning Workshop*. Technical Report. University of Central Florida and Office of Naval Research.

[37] Y. Zhao, S. Szpiro, and S. Azenkot. 2015. ForeSee: A Customizable Head-Mounted Vision Enhancement System for People with Low Vision. *Proceedings of the International ACM SIGACCESS Conference on Computers & Accessibility* (2015), 239–249.