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.2022.3150503

The Effect of Context Switching, Focal Switching Distance, Binocular and Monocular Viewing, and Transient Focal Blur on Human Performance in Optical See-Through Augmented Reality

Mohammed Safayet Arefin, *Student Member, IEEE*, Nate Phillips, *Student Member, IEEE*, Alexander Plopski, Joseph L. Gabbard, *Senior Member, IEEE*, and J. Edward Swan II, *Senior Member, IEEE*

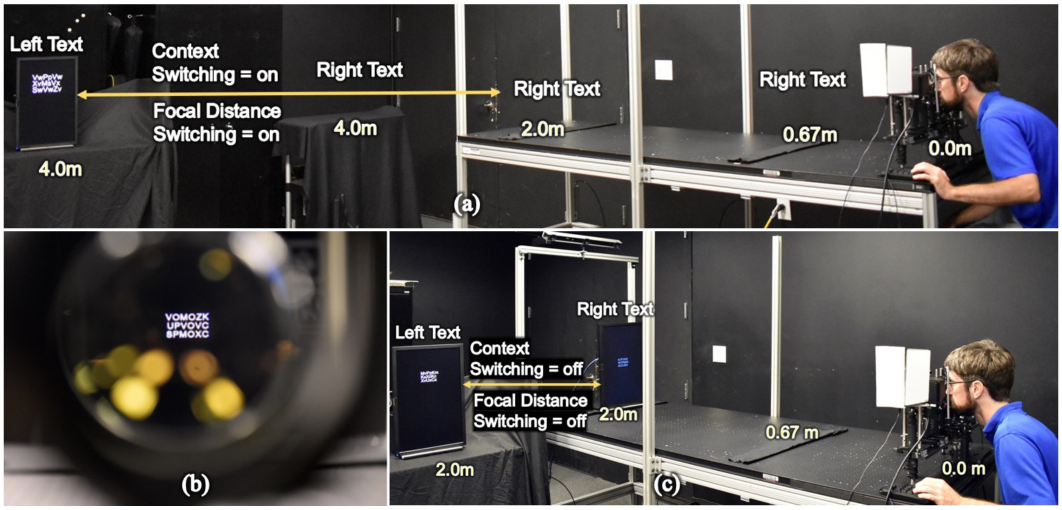

Fig. 1. Participants performed a text searching task, in an experiment that switched contexts between the real world and augmented reality (AR), at either matched or mismatched focal distances. (a) A participant observing the left text on a monitor at 4 meters distance, and the right text in AR at one of three focal distances: 4, 2, or .67 meters. (b) View of the right text through the custom-built AR Haploscope. (c) The participant observing both the left text and the right text on a monitor at 2 meters distance.

Abstract— In optical see-through augmented reality (AR), information is often distributed between real and virtual contexts, and often appears at different distances from the user. To integrate information, users must repeatedly switch context and change focal distance. If the user's task is conducted under time pressure, they may attempt to integrate information while their eye is still changing focal distance, a phenomenon we term *transient focal blur*. Previously, Gabbard, Mehra, and Swan (2018) examined these issues, using a text-based visual search task on a one-eye optical see-through AR display. This paper reports an experiment that partially replicates and extends this task on a custom-built AR Haploscope. The experiment examined the effects of context switching, focal switching distance, binocular and monocular viewing, and transient focal blur on task performance and eye fatigue. Context switching increased eye fatigue but did not decrease performance. Increasing focal switching distance increased eye fatigue and decreased performance. Monocular viewing also increased eye fatigue and decreased performance. The transient focal blur effect resulted in additional performance decrements, and is an addition to knowledge about AR user interface design issues.

Index Terms—Augmented reality, context switching, focal distance switching, transient focal blur, accommodation

1 Introduction

Optical see-through (OST) augmented reality (AR) superimposes computer-generated virtual information on a user's view of the real world, usually presented through a head-mounted display (HMD). Often, information relevant to the user's task is distributed between real and virtual contexts and appears at different focal distances from the user. Therefore, to integrate the information, the user must repeatedly switch context and refocus the eyes. Here, *context switching* refers to switching visual and cognitive attention between real and virtual information, while *focal distance switching* refers to accommodating the eyes' lenses to see, in sharp focus, information at a new focal distance. *Focal switching distance* is the distance over which the lenses must be accommodated. Changing accommodation to a new focal distance can take as long as 425 milliseconds [6], and during this time period, information will appear blurry. Here, we refer to this blur as *transient focal blur*.

These phenomena arise in many OST AR application use cases. For example, consider a car manufacturer using an OST AR display to assist in assembly. A virtual text label is applied to a car part, and the user's task requires them to see both the label and the part. If the virtual text label is presented at the same distance as the part, then there is context switching, but no focal distance switching. If the text label

- *Mohammed Safayet Arefin, Nate Phillips and J. Edward Swan II are with Mississippi State University, USA. E-mail: arefin@acm.org, Nathaniel.C.Phillips@ieee.org, swan@acm.org*
- *Alexander Plopski is with University of Otago, New Zealand. E-mail: alexander.plopski@otago.ac.nz*
- *Joseph L. Gabbard is with Virginia Tech, USA. E-mail: jgabbard@vt.edu*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

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.2022.3150503

is presented at a different distance than the part (e.g., at the location where the part must be placed), then there is both context switching and focal distance switching. Finally, if there is no virtual text label for a given part, but both the part and the location where it must be placed have printed labels that must be compared, then there is focal distance switching, but no context switching.

Previous research has found that context switching can reduce performance, both for general information displays [52] and when context switching between AR and real world content [16, 24]. In addition, frequent focal distance switching can result in excessive eye strain, visual fatigue, and reduced task performance [3, 4, 13, 15, 23, 26, 39]. Of these papers, only Hoffman et al. [23] used a custom-built laboratory display. All of the others employed commercial, off-the-shelf OST AR displays, with inherent limitations in the consistency of presented depth cues. For example, Wang Baldonado [52] used standard computer monitors, Huckauf et al. [24] and Gabbard et al. [16] used monitors and an one-eye Microvision Nomad, Neveu et al. [39] used a television and a Sony Glasstron, Imamov et al. [26] used an HTC Vive Pro, Eiberger et al. [15] used an Epson Moverio BT-100, and Drouot et al. [13] used a Microsoft Hololens 2.

While papers using off-the-shelf displays serve as important foundations for examining the phenomena of context and focal distance switching, to fully understand why performance decrements are observed, a *vision science approach* is needed (e.g., [9, 20, 23, 28, 33]). These and related papers have inspired two aspects of the approach reported here: (1) They generally use *custom laboratory-built displays*, which allow precise control over all relevant optical and visual parameters. And (2), the experiments often include a *monocular condition*, where the non-dominant eye is covered. The monocular condition is motivated by the importance of stereo vision for many human tasks, and the related depth cues of stereo disparity and ocular vergence [11]. Including a monocular condition therefore allows stereo vision effects to be separated from other effects, and helps explain experimental findings in the context of the human visual system. Therefore, a unique contribution of this paper is that the phenomena of interest were carefully examined using a custom laboratory-built, OST AR display, and under conditions of both binocular and monocular viewing. This not only allows replicating findings seen in previous studies using commercial off-the-shelf displays, but critically allows the findings to be attributed to specific elements of the human visual system. Such results in turn could be used to inform future AR hardware design and practitioners' selection of AR hardware features, when considering use cases where focal and context switching demands can be predicted in advance.

Accordingly, the purpose of the current experiment was to systematically investigate, in OST AR, the phenomena of *context switching, focal switching distance, monocular and binocular viewing*, and *transient focal blur*. This was accomplished by a partial replication[1] and extension of the task and experiment reported by Gabbard et al. [16], on a custom-built optical testbed designed specifically to examine these issues, i.e., an AR Haploscope (Figs. 1 and 3).

Gabbard et al. [16] expressed the concern that their findings might be specific to the Nomad AR display, which used a unique display technology. In addition, their experimental design did not fully cross the conditions of context switching and focal distance switching, and therefore could not fully consider how these conditions interact. Moreover, their experiment used binocular vision while wearing an AR display that only covered one eye. This condition, which we term *semi-binocular* viewing in this paper, matches expected use cases of one-eye AR displays in task domains such as order picking [47]. However, in the context of the vision science approach advocated here, their experiment did not fully test either binocular or monocular viewing of virtual content. The experiment reported here addresses all of these concerns.[2]

[1] A *partial replication* is a replication of an original experiment with intentional but small modifications [21].

[2] Portions of this work are reported in a poster abstract [2] and an MS thesis [1].

2 Background and Related Work

2.1 Accommodation and Vergence

Accommodation is the ability of the eye to observe an object in sharp focus [53]. The primary stimulus that drives the accommodative response is a *blur gradient*, where blur is reduced as the eye adjusts its focal length. No current commercial augmented or virtual reality displays produce a blur gradient that drives an accommodative response. However, Cholewiak et al. [9] developed a laboratory-built display that does drive an accommodative response; their display incorporates both blur and chromatic aberration.

Along with accommodation, viewing an object requires rotational *vergence eye movements*. When fixating on an object closer than where the eyes were previously verged, the eyes *converge* and rotate towards each other, while when fixating on an object farther than where the eyes were previously verged, the eyes *diverge* and rotate away from each other. The primary stimulus that drives the vergence response is *stereo disparity*; when fixating on an object the eyes verge until the images can be fused into a single image [30]. In addition to these individual stimuli to accommodate and verge, both responses are linked to each other, so that accommodation drives vergence (*accommodative vergence*) and vergence drives accommodation (*vergence accommodation*). Pupil diameter is also involved [30, 36]. The link between them is known as the *vergence-accommodation reflex*. Although when viewing objects in the real world, the stimuli to accommodate and verge co-vary in depth, the human visual system can override the vergence-accommodation reflex. This is often necessary when viewing information on a stereo display. The accommodative demand of a display is either the physical distance of the display, or—as in an OST AR display—the focal distance of the display's optical system. When the stereo disparity of a virtual object matches this distance, the visual system does not have to override the vergence-accommodation reflex. However, often the stereo disparity of a virtual object is at a different distance, and then the visual system does have to override the reflex. This causes an *accommodation-vergence mismatch*. This mismatch is the source many perceptual problems for stereo displays, including distortions in perceived depth and size [31], eye strain [29], double-vision [37], reduced user performance, and increased cognitive load [23, 32].

To date, commercial OST AR displays have generally presented virtual objects at a single fixed focal distance; recent examples include the Microsoft HoloLens (versions 1 and 2) and the Google Glass. However, the Magic Leap One presents virtual objects at two fixed focal distances [46]. To the best of the authors' knowledge, the Microvision Nomad, used in Gabbard et al. [16] and other studies [24, 47], is the only commercial display with an adjustable focal knob. Some laboratory-built displays have used lenses that can be refocused by software [9, 28]. However, even for displays with an adjustable focus, only objects at the current focal depth can be seen without accommodation-vergence mismatch.

Although the eyes are capable of very rapid *saccadic* eye movements, changing accommodation and vergence to fixate on an object at a new distance is relatively slow [30]. Up to age 20, the human eye requires 360 milliseconds to accommodate from far to near and 380 milliseconds to accommodate from near to far. After the age of 20, the time required to accommodate from near to far remains relatively constant, but the time required to accommodate from far to near increases [6, 22]. The properties of the visual stimulus can also change the time required to accommodate, and can be as long as 425 milliseconds [7].

2.2 Visual Fatigue

Lambooij et al. [32] defined visual fatigue as "physiological strain or stress resulting from excessive exertion of the visual system". Visual fatigue has a wide range of visual symptoms, including eye strain, blurred vision, difficulty focusing, ache around the eyes, soreness around the eyes, and others [50]. Many activities can cause visual fatigue; some examples include using a computer screen for many hours, reading under inadequate lighting, and reading poorly printed text. However, most relevant here is that frequently changing accommodation and vergence leads to eye fatigue, eye strain, and reduced task perfor-

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.2022.3150503

mance [15, 23, 35, 39, 40, 50]. In the experiment partially replicated here, Gabbard et al. [16] found significant effects of eye fatigue.

Visual fatigue is a particular concern for task domains where an AR display would be used daily for many hours. For example, Schwerdtfeger et al. [47], in an early study (mid-2000's) of using an AR system for an industrial order picking task, which involved repeated context and focal distance switching, found that after only 2 hours, 50% of the participants reported focus problems, some significant enough that they had to pause the task.

2.3 Context Switching and Focal Distance Switching

Although switching of context and focal distance in AR is frequent, little research has considered the impact on human performance. Gabbard et al. [16] first explored the interaction between context and focal distance switching. They used a text-based visual search task that required participants to integrate information from both real and AR environments. For displaying AR information, they used a Microvision Nomad see-through AR display, a one-eye display that uses a laser-based retinal scanning technology [51]. Unique among commercially-available displays, this allowed the display to change focal distances. Their study found that context switching had a negative impact on performance when information was presented at 6 meters, but not at closer distances of 2 or .7 meters. However, context switching resulted in greater eye fatigue at all three distances. Focal distance switching resulted in reduced performance, and additional performance reductions were attributed to transient focal blur experienced while switching focal distances.

Eiberger et al. [15] examined the combined effects of context and focal distance switching. They were motivated by the application of integrating information between a smartwatch and more distant environmental surfaces. They simulated environmental viewing with an Epson Moverio BT-100 display, which presents collimated imagery (infinite focal distance) at a stereo disparity of 3.7 meters. This was compared to a projected image at .3 meters (a typical smartwatch distance). They used a graphical visual search task. Context and focal distance switching resulted in a higher task completion time and a larger error rate. Recently, Drouot et al. [13] also examined the combined effect of context and focal distance switching, using a Microsoft Hololens 2 at 1.5 or 2.0 meters distances, and a graphical visual search and target detection task. They found a negative performance effect for distance switching, but no effect for context switching.

Two previous studies have examined AR context switching. The first was Huckauf et al. [24], who used semi-binocular viewing of a one-eye Microvision Nomad display focused at .61 meters, and a monitor placed at the same distance. They found that context switching between the displays reduced performance on several different visual tasks. Most recently, Imamov et al. [26] investigated the issue of context switching by displaying information on two interfaces within a VR environment (simulating AR interaction). Their research found that context switching increased task completion time and decreased user comfort.

3 Method

The purpose of the current experiment was to systematically investigate, in OST AR, the phenomena of *context switching, focal switching distance, binocular and monocular viewing*, and *transient focal blur*. Several hypothesis were developed:

H1: Context switching would reduce performance and increase eye fatigue.

H2: Larger focal switching distances would reduce performance and increase eye fatigue.

H3: The performance reduction during focal distance switching, attributed to the transient focal blur effect in Gabbard et al. [16], world replicate under different conditions of context switching and viewing condition.

Task Description	Left text	Right text
Participants identified the doubled target letter 'O' in the left text, and then counted the number of target occurrences in the right text. Here, the correct answer is '1'.	sKvKuS mUpKuP sOoMsP	POXCSK SZSXMM VKUKPC

Fig. 2. The experimental subtask. In Gabbard et al. [16], the left text was presented in either AR or the real world, while in the current experiment, the right text was presented in either AR or the real world (Fig. 1).

3.1 Experimental Subtask and Task

A text-based visual search task was employed. The task replicates the one used by Gabbard et al. [16], and is based on previous versions of the task [12, 17, 18]. The task was motivated by reading AR text labels and comparing them to text in the real world, similar to the warehouse logistics task described by Schwerdtfeger et al. [47]. The task does not examine the semantic interpretation of what is read, but instead measures the low-level visual identification of letters. It has the important property that successful completion requires integrating information from two different blocks of text. When one of these text blocks is presented in AR and the other in the real world, the task requires context switching. Alternatively, when both text blocks are presented in the real world, the task does not require context switching. In addition, the text blocks can be presented at different focal distances, or the same focal distance. The absolute difference between two focal distances provides the amount of focal switching distance. Therefore, the task affords testing every 2×2 combination of context switching and focal distance switching.

Participants observed two side-by-side text blocks, the left text and the right text (Fig. 2). Each text block comprised three text strings, and each text string comprised six letters. The strings were drawn from a 24-character alphabet, which consisted of the standard 26-character English alphabet without the letters i, j, and l. The rationale for removing these letters is that in the employed sans-serif font, upper- and lower-case versions of these letters were too difficult to distinguish from each other. The left text consisted of pairs of letters, alternating between upper and lower case, while the right text was all upper-case letters.

The task consisted of a series of subtasks. Each subtask required three actions: (1) Searching the left text for the target letter, which was encoded by a pair of side-by-side identical letters, one upper case and the other lower case (e.g., "Oo" in Fig. 2). In the left text there was always exactly one target letter. (2) Searching for the target letter in the right text. The target letter could appear at most once in each line of text, and could appear in total 0, 1, 2, or 3 times. Therefore, if the target letter appeared three times, it appeared in all three lines of text, and if the target letter appeared two times, it appeared in two lines of text. In Fig. 2, the target letter "O" appears once, in the first line of text. (3) After counting the number of target letters in the right text, the target letter count was entered on a numeric keypad (Fig. 3). Participants knew that the correct answer was one of 0, 1, 2, or 3, and they knew that each line could contain at most one target letter.

The task was to complete 5 subtasks within 25 seconds. As described by Gabbard et al. [16], when the experimental task was first developed, a group of pilot participants took on average 30 seconds to complete 5 subtasks. Therefore, reducing the time limit to 25 seconds introduced performance pressure, in order to better differentiate experimental conditions. At the beginning of a task, the left text and right text appeared. When a participant entered their answer for the first subtask, the right text remained, but a new left text appeared, giving a new target letter. Each subtask resulted in a new left text, but the right text remained for the entire task. The task ended when the first of two events occurred: the participant completed 5 subtasks, or 25 seconds had elapsed. When the task ended, both text blocks disappeared, and no additional input from the keypad was accepted.

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.2022.3150503

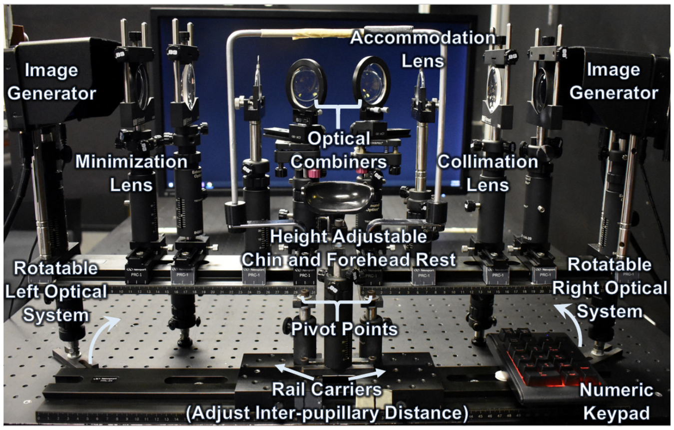

Fig. 3. The Augmented Reality (AR) Haploscope allows the precise adjustment of focal distance and vergence angle.

3.2 Apparatus

A custom-made *augmented reality haploscope*—an AR display mounted on an optical workbench—was used (Figs. 1 and 3). The AR haploscope was based on the design presented by Singh et al. [48], and further described by Phillips et al. [42, 43]. As shown in Fig. 3, the haploscope consisted of left and right optical systems, which rotated around pivot points that were approximately aligned with the rotational axes of the participant's eyes. Additional adjustments fit the participant's inter-pupillary distance and face height. Each optical system was composed of a fixed minimization lens, a fixed collimation lens, and an adjustable accommodation lens. The optics projected the images from two small, high-resolution image generators to varying focal distances. The image generators, Feelworld F570 5.7" 4K, had a diagonal size of 14.5 cm, a display resolution of 1920×1080 pixels, a maximum luminance of 460 cd/m², and a contrast ratio of 1400:1. A set of optical combiners, with 15% reflectivity, then combined these images with the view of the real world.

For this experiment, the AR haploscope afforded precise adjustment of focal distance and vergence angle. Focal distance was adjusted by changing the power of the accommodation lens. Because different lens powers also change the magnification of the optical system, the image size, as presented by the image generators, was carefully calibrated to take this magnification change into account [43]. Vergence angle was adjusted by rotating the left and right optical systems around the pivot points. The rotation amount depended on both the focal distance and the participant's inter-pupillary distance. The final rotation angle was calculated by the experimental software on a per-participant, per-focal demand basis, and was tracked by an optical tracking system (OptiTrack-V120 Trio).

The haploscope could be used in both an AR mode, presenting virtual right text (Fig. 1*a*, *b*), as well as a real mode, where participants looked through the optical combiners at the right text presented on a monitor (Fig. 1*c*). During real mode operation, the image generators were switched off, and therefore the combiners added no additional luminance to the participant's eyes. Two standard PC monitors were also used, Dell U2211H, with a diagonal size of 55 cm, resolution of 1920×1080 pixels, and a maximum luminance of 23.23 lumens (Fig. 1*c*). In addition, an illuminated numeric keypad (Fig. 3) collected participant responses.

3.3 Setup

The haploscope was mounted on the end of a 244 cm by 92 cm optical breadboard (Fig. 1), which was supported by a custom-built aluminum table. Four of the table legs extended above the breadboard surface and supported the optical tracking system. Participants sat on a tall height-adjustable office chair, which allowed participants of different heights to comfortably look through the haploscope. The laboratory setting (Fig. 1*a*) allowed viewing real world objects at distances of up to 4 meters. The black walls were both undistracting and light absorbing.

A sans-serif font (Arial) was used, based on common usage for text labels, as well as previous work that demonstrated good readability on computer displays [5]. The text was rendered in white to maximize its luminance, and its height was adjusted so that capital letters spanned a visual angle of 22 arcminutes, meeting Federal Aviation Administration recommendations for text legibility [54]. The visual angle of text size was constant at each distance. This differed from Gabbard et al. [16], where the text size varied with distance. While in their study relative size was a depth cue [11], the current experiment prioritized text legibility over this cue.

The right text was observed by looking straight ahead (Fig. 1), which allowed the right text to be viewed in either AR or on a monitor. The left text was always viewed on a monitor. The left text monitor was positioned so that, regardless of text distance, glancing between the two text blocks always required an eye rotation of 26.2 degrees.

The experimental control program ran on an Alienware Windows 10 desktop computer. It was written in C++ and Perl, and used the OpenCV library.

3.4 Independent Variables

The experimental variables were *context switching*, *focal switching distance*, and *viewing*. Of these, focal switching distance was a function of *reference distance* and *test distance* (Fig. 1, Table 1).

Context Switching (*no, yes*): When participants viewed both the left and right texts on monitors (Fig. 1*c*), *context switching* was not required to complete the task. However, when participants viewed the right text in AR (Fig. 1*a*), context switching was required.

Focal Switching Distance (0, 1.33, 2, 3.33 meters): The *focal switching distance* was a function of the reference and test distances. The *reference distance* was the distance from the participant to the left text, always viewed on a monitor, while the *test distance* was the distance from the participant to the right text, viewed in either AR or a monitor (Fig. 1). When viewing real text, monitors were placed .67, 2, or 4 meters from the participant's eye position. When viewing AR text, the AR haploscope encoded the distance by using an accommodation lens of the correct power (1.5, .5, or .25 diopters), and by setting the vergence angle according to the distance and the participant's inter-pupillary distance (Fig. 3). The *focal switching distance* was computed as $f = | r - t |$, where f is the focal switching distance, r the reference distance, and t the test distance.

Viewing (*monocular, binocular*): In the *binocular* condition, participants performed the experiment with both eyes open. In the *monocular* condition, participants covered their non-dominant eye with an eye patch, and performed the experiment with their dominant eye.

Repetition (1 to 5): Each combination of viewing, context switching, reference distance, and test distance was repeated five times.

3.5 Dependent Variables

Five dependent variables were measured or calculated: *number of subtasks completed, number of subtasks correct, undercount errors, overcount errors,* and *eye fatigue*. The same dependent variables were collected by Gabbard et al. [16].

Number of Subtasks Completed (0 to 5): As discussed in Section 3.1 above, each task involved completing up to 5 subtasks within 25 seconds. The *number of subtasks completed* (also, *subtask completion*) served as a primary performance measure.

Number of Subtasks Correct (0 to 5): For each subtask, the error was calculated as *error = participant target count – correct target count*, where each target count ranged from 0 to 3. When *error* = 0, the subtask was correct. The *number of subtasks correct* (also, *subtask accuracy*) served as a complementary performance measure to the number of subtasks completed. Note that, for any set of experimental conditions, *number of subtasks correct* $\leq$ *number of subtasks completed.*

Undercount and Overcount Error (–3 to +3): When $error \neq 0$, the subtask was not correct. $Error > 0$ indicated an *overcount error*, where a participant counted more letters than were displayed. Overcount errors

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.2022.3150503

Table 1. Experimental Design. Viewing is encoded by letter: m (monocular), b (binocular).

Real World to Real World Conditions (*Context Switching* = no)						
	Test Distance (R):					
Reference Dist. (R):	R1 (0.67)	R2 (2.0)	R3 (4.0)	R1 (0.67)	R2 (2.0)	R3 (4.0)
R1 (0.67)	mR1R1	mR1R2	mR1R3	bR1R1	bR1R2	bR1R3
R2 (2.0)	mR2R1	mR2R2	mR2R3	bR2R1	bR2R2	bR2R3
R3 (4.0)	mR3R1	mR3R2	mR3R3	bR3R1	bR3R2	bR3R3
Real World to AR Conditions (*Context Switching* = yes)						
	Test Distance (A):					
Reference Dist. (R):	A1 (0.67)	A2 (2.0)	A3 (4.0)	A1 (0.67)	A2 (2.0)	A3 (4.0)
R1 (0.67)	mR1A1	mR1A2	mR1A3	bR1A1	bR1A2	bR1A3
R2 (2.0)	mR2A1	mR2A2	mR2A3	bR2A1	bR2A2	bR2A3
R3 (4.0)	mR3A1	mR3A2	mR3A3	bR3A1	bR3A2	bR3A3

ranged from 1 to 3. *Error* < 0 indicated an *undercount error*, where a participant missed counting letters that were displayed. Undercount errors ranged from –1 to –3.

Eye Fatigue (1 to 7): After completing each block of five task repetitions, participants were asked to rate their eye fatigue by answering the question “Please rate the condition of your eyes”. A 7 point bi-polar rating scale, ranging from “very rested” to “very fatigued”, was displayed on the physical monitor. Participants responded by pressing the appropriate number key on the numeric keypad.

3.6 Experimental Design and Counterbalancing

Experimental Design: The experimental design is shown in Table 1. The upper half shows the real world to real world conditions, where no context switching occurred, while the lower half shows the real world to AR conditions, where context switching occurred. The left half shows conditions viewed monocularly, while the right half shows conditions viewed binocularly. Within the resulting 2 (*context switching*) $\times$ 2 (*viewing*) design, all 9 combinations of the 3 reference and 3 test distances were presented. The cells where the reference and test distances match are highlighted; these are cells where focal distance switching was not required. The cells in Table 1 uniquely label each combination of conditions. For example, cell mR2R3 indicates that participants monocularly viewed the reference text in the real world at a distance of 2 meters, and the test text in the real world at a distance of 4 meters. Here, when looking between the left and right text, the task did not require switching context, but did require switching focal distance by 2 meters. For another example, cell bR1A1 indicates that participants binocularly viewed the reference text in the real world at a distance of .67 meters, and the test text in AR, also at a distance of .67 meters. Here, when looking between the left and right text, the task required switching context, but did not require focal distance switching.

Comparison to Gabbard et al. [16]: Table 1 facilitates comparing the design of the current experiment to the design of Gabbard et al. [16]; see Tables 1 and 2 in Gabbard et al. In the real world to real world conditions, with no context switching, they only studied matched distances (the shaded cells in Table 1 [16]). In addition, in Gabbard et al. participants viewed real world text binocularly, and AR text semi-binocularly. Finally, relative size and motion parallax were used to encode a virtual distance to AR text, where this virtual distance varied independently from focal distance (Table 2 [16]). Gabbard et al. made these decisions for sound experimental reasons: they studied a one-eye AR display, where the expected use case is looking at real world content binocularly, and AR content semi-binocularly. And they reduced the size of the experimental design to the point where participants could complete the experiment in a single session of 2 hours or less. In contrast, the current experimental design expands and fully counterbalances the design (Table 1), but required participants to attend two separate experimental sessions of 2 hours or less.

Counterbalancing: As each participant experienced all 36 conditions shown in Table 1, a within-subjects experimental design was used. Between participants, the presentation order of viewing was counterbalanced; half the participants experienced the monocular condition followed by the binocular condition, while the remaining experienced the opposite order. The two levels of viewing were experienced in two separate experimental sessions. For 15 of the 24 participants, 2 to 3 days elapsed between these sessions. Because of scheduling complexities, for the remaining participants several weeks elapsed; the longest time period was 30 days. Also between participants, within each session the presentation order of context switching was counterbalanced; half the participants experienced context switching in the order *no*, *yes*, while the other half experienced the order *yes*, *no*. Therefore, with two levels of viewing and two levels of context switching, there were four possible condition orderings for each participant. A 4$\times$4 Latin square controlled the presentation order for each participant, and therefore the presentation order was fully counterbalanced for each group of four participants. Within each participant, the presentation order of the remaining independent variables, reference distance and test distance, was randomly permuted. Therefore, each participant completed 2 (*viewing*) $\times$ 2 (*context switching*) $\times$ 3 (*reference distance*) $\times$ 3 (*test distance*) $\times$ 5 (*repetition*) = 180 tasks, where each task comprised as many as 5 subtasks.

3.7 Participants

24 participants from the Mississippi State University community participated in this experiment; 12 were male and 12 female. The mean age of the participants was 22.9 years; age ranged from 18 to 31. The participants had a mean inter-pupillary distance of 63.1 mm; 17 participants were right-eye dominant (71%), and 7 left-eye dominant (29%), which agrees with the expected distribution of eye dominance [44]. No vision corrective restriction was provided to filter the participants; 13 participants wore corrective lenses, while 11 did not require correction. Participants were young enough to not exhibit presbyopia [14]. 13 participants were compensated with course credit, and 11 were compensated at a rate of 12 USD per hour.

3.8 Procedure

Participants attended two experimental sessions. Each session lasted $\sim$1.5 hours, within a range of 1 hour for the fastest participants to 2 hours for the slowest.

Pre-Trial Tasks, First Experimental Session: At the beginning of a participant's first session, they received a short explanation of the experiment and related procedures, and they filled out a consent form and a pre-experiment questionnaire. Next, a commercial pupilometer was used to measure the participant's inter-pupillary distance at optical infinity. Then, the participant's dominant eye was determined with the Miles [34] test. The experimental task and subtask were then described. Using a paper version of the task, the participant next performed 2, 3, or more task repetitions. Each task was comprised of subtasks that covered all of the possible number of times that the target letter could appear in the right text (0, 1, 2, or 3). The experimenter continued these practice trials until they were convinced that the participant thoroughly understood the subtask and task.

Pre-Trial Tasks, Second Experimental Session: At the beginning of a participant's second session, they received a short reminder of the experiment and related procedures. They then practiced the paper version of the task several times, until the experimenter was convinced that the participant again remembered and understood the subtask and task.

Apparatus Adjustments: Next, the haploscope's rail carriers were adjusted so the distance between the pivot points matched the participant's inter-pupillary distance (Fig. 3) [43]. The chin and forehead rest were adjusted, as well as the height of the office chair and the position of the numeric keypad, so the participant could sit comfortably and look through the haploscope (Fig. 1). Once calibrated, the haploscope pivot points were approximately located directly under the rotational center of the participant's eyes. The participant's inter-pupillary distance was entered into the experimental control software. For every AR test distance, to present the correct vergence angle, the control software calculated the required angle of the left and right optical systems. Next,

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.2022.3150503

if the current experimental session used monocular viewing, the participant's non-dominant eye was covered with an eye patch, and the image generator on the non-dominant eye side of the haploscope was turned off.

Experimental Trials: Participants next completed the tasks for their experimental session, which comprised either all of the monocular or all of the binocular conditions shown in Table 1. Each condition was tested with a block of 5 task repetitions.

Before beginning a new condition, the experimenter ensured that the participant could see every letter of the left and right text by performing a calibration process. During this process, a left and right text was shown; neither was a text that would be used in the subsequent block of trials. The experimenter asked the participant if they could see every letter of both texts. Because the haploscope optics have a relatively narrow field of view, the tolerance for the position of the participant's eyes was tight, and occasionally one or both texts would appear off-center, and part of the text would to be hidden from one or both eyes. Using key presses, the experimental control software allowed the experimenter to quickly shift the position of both the left and right text on the monitor, or AR image generator, by a small amount up, down, left, or right. If needed, the experimenter adjusted these positions until the participant indicated that both texts were centered, and they could see every letter. During binocular viewing, the experimenter had the participant close one eye and then the other, and performed this calibration until the participant could see every letter with each eye, thus ensuring that all letters were seen with binocular vision.

The participant then completed 5 task repetitions for the condition. After completing the last repetition, the participant was prompted to report their eye fatigue. After this, the experimenter asked the participant to relax, sit back in their chair, and close their eyes. During this rest period, the experimenter changed the physical apparatus as needed for the next condition: either moving monitors for the right and / or left text, or adjusting the accommodation lenses and vergence angle of the haploscope.

Post-Trial Tasks: After completing all the experimental trials, the experimenter conducted an informal interview to gather additional information and insights from the participant, and then thanked the participant and dismissed them. The experimenter then typed up contemporaneous notes about the experimental session.

4 Results and Discussion

4.1 Analysis

When one of the independent variables was continuous, data was analyzed by examining the slopes and intercepts of linear equations, and multiple regression was used to determine if the slopes and intercepts significantly differed [10, 41]. When all independent variables were categorical, data was analyzed by examining means and standard errors, and significance was determined by repeated-measures factorial ANOVA. Through an arrangement with the authors of Gabbard et al. [16], the data from that experiment was available to the authors, and the analysis sometimes compares data from both experiments.

Results are generally shown in scatter plots such as Fig. 4. Here, the x axis, the *continuous independent variable*, gives the distance to the reference text, and the y axis, the *dependent variable*, shows the number of subtasks completed (upper row) and correct (lower row). The grey points are the number of subtasks completed for each (x, y) value. Context switching, the *categorical independent variable*, is indicated by the color and position of violin plots, which summarize the point distributions for each level of context switching. As indicated by the caption, each panel displays 144 grey points, but with substantial overlap.

The multiple linear regression procedure, from Pedhazur [41] chapter 12, fits one or two linear regression lines in each panel of each scatter plot. The procedure is separately applied to each panel. The F-tests generated by the procedure are given in tables located in the Appendix (see supplemental files). Each graph has a corresponding table; Table A2 corresponds to Fig. 4. Each graph panel has a corresponding row in the associated table; note panels a to f in Fig. 4, and rows a to f in Table A2.

The multiple linear regression analysis proceeds in four steps:

1. Two linear regressions are generated, one for each level of the categorical independent variable. An F-test then determines if the slopes significantly differ. If they do, as in Fig. 4*a* (Table A2*a*: *slope diff*), both linear regressions are reported as the best overall description of the data in the panel. Two lines are drawn, and two linear equations given. The interaction between the continuous and categorical independent variables is significant.

2. If the slopes do not differ, then the slopes are set to a common value, and an F-test determines if the intercepts significantly differ. If they do, as in Fig. 5*a* (Table A3*a*: *intercept diff*), these two linear regressions are reported as the best overall description. Two lines are drawn, with a common slope, and two linear equations given. The main effect of the categorical variable is significant.

3. If the intercepts do not differ, as in Fig. 4*b* (Table A2*b*), then a single linear regression is reported as the best overall description. One line is drawn, and one linear equation is given.

4. If the two slopes do not differ, an additional F-test determines if the single slope differs from 0. This can either be the common slope of two regressions, as in Fig. 5*a* (Table A3*a*: *slope 0*, same degrees of freedom as the intercept test), or the slope of a single regression, as in Fig. 4*b* (Table A2*b*: *slope 0*, one degree of freedom larger than the intercept test). If the slope differs from 0, then the main effect of the continuous independent variable is significant.

When applying these F-tests, to properly account for repeated measurements, responses are averaged over repetition per participant, per experimental cell (Table 1) [41]. For example, each panel in Fig. 4 shows 144 data points, which are averaged from 720 data points for the Haploscope (5 repetitions), or 576 data points for the Nomad (4 repetitions).

The multiple regression analysis yields two measures of effect size: (1) R^2, the overall percentage of variation explained by the linear model, and (2) dR^2, the percentage of variation explained by the categorical variable. Both R^2 and dR^2 are reported for every panel. In addition, if two linear regressions are reported, then d, the distance between the lines in y axis units, are reported. If the slopes differ (e.g., Fig. 4*a*), signed distances are reported for the leftmost and rightmost data points along the x axis (for Fig. 4*a*, $x = .7$ and 6 meters). If the slopes do not differ (Fig. 5*a*), an unsigned distance is reported. Sometimes the value of the slope, b, is also discussed.

4.2 Context Switching

Context switching was expected to reduce task performance and increase fatigue (**H1**). Context switching was examined by comparing cells where context switching occurred, but focal distance was held constant: the shaded cells in Table 1 were compared between the conditions of context switching = *no* and context switching = *yes*. These cells contain 30% of the collected data.

4.2.1 Task Performance

The task performance effects of context switching and reference distance are analyzed in Fig. 4. The left-hand column shows the relevant data from Gabbard et al. [16] (display = *Nomad*), under the *semi-binocular* viewing condition. The center and right columns show the data from the current experiment (display = *Haploscope*), under the viewing conditions of *monocular* and *binocular*. The upper row shows performance in terms of *subtasks completed*, while the lower row shows performance in terms of *subtasks correct*.

On the Nomad, there was a significant interaction between context switching and reference distance (panels a, d): at short distances of .7 and 2 meters, context switching had very little effect, but at the longer distance of 6 meters, $d = 1.043$ (1.088) fewer subtasks were completed

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.2022.3150503

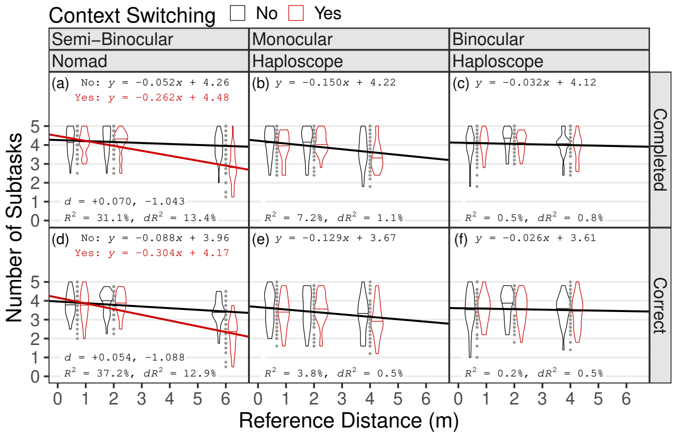

Fig. 4. When examining data matched in depth, on the Nomad display (Gabbard et al. [16]) context switching reduced performance at the far reference distance. However, on the Haploscope context switching had no effect, and increasing reference distance only reduced performance under monocular viewing. *Nomad*: Data from highlighted cells in Tables 1 and 2 [16]: R1R1, R2R2, R3R3; V1F1R1, V2F2R2, V3F3R3. *Haploscope*: Data from highlighted cells in Table 1: mR1R1, mR2R2, mR3R3; mR1A1, mR2A2, mR3A3; bR1R1, bR2R2, bR3R3; bR1A1, bR2A2, bR3A3. Table A2 (Appendix; see supplemental files) shows the related *F*-tests. Each panel displays 144 data points.

(correct). The effect is strong, explaining $dR^2 = 13.4\%$ (12.9%) of the variation. In the current experiment, this effect was not repeated; there was no effect of context switching with either monocular or binocular viewing. However, there was an effect of reference distance: under monocular viewing, increasing reference distance resulted in reduced performance, at a rate of $b = .150$ $(.129)$ subtasks completed (correct) per meter. While significant, this effect only explains $R^2 = 7.2\%$ (3.8%) of the variation, much less than what is explained for the Nomad. Under binocular viewing, there was no effect of either context switching or reference distance.

4.2.2 Eye Fatigue

The eye fatigue effects of context switching and reference distance are analyzed in Fig. 5. Other than the y axis displaying fatigue, the graph structure is the same as Fig. 4. On the Nomad, there was a significant main effect of context switching on fatigue, but no effect of reference distance: context switching increased eye fatigue by $d = 1.125$ units at all distances. On the Haploscope, this effect was repeated. Under monocular viewing, context switching significantly increased eye fatigue by $d = .875$ units at all distances. In addition, there was a marginally significant main effect of reference distance ($p = .08$), where increasing distance resulted in increased fatigue, at a rate of $b = .149$ units per meter. Under binocular viewing, context switching also significantly increased eye fatigue by $d = .486$ units, with no effect of reference distance. Therefore, context switching increased eye fatigue across all conditions, but with different magnitudes in each condition.

4.2.3 Discussion

It was hypothesized that context switching would decrease performance and increase eye fatigue (**H1**). Figs. 4 and 5 both directly compare the previous findings with those of the current experiment. On the Haploscope, context switching had no effect on task performance, but it did increase fatigue. Therefore, the current results partially support hypothesis **H1**.

As described by Gabbard et al. [16], on the Nomad the reason for reduced performance was blurry vision, especially at the far distance of 6 meters. Likely reasons for this blurry vision were distance from the resting point of accommodation (for most participants less than 50 centimeters [25]), laser speckle in the Nomad display, and a smaller font

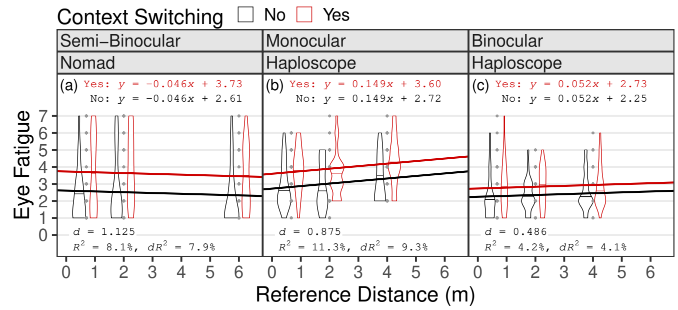

Fig. 5. When examining data matched in depth, context switching increased eye fatigue for both the Nomad display (Gabbard et al. [16]) and the Haploscope. On the Haploscope, the amount of increased eye fatigue was higher with monocular viewing, compared to binocular viewing. Also under monocular viewing, increased reference distance resulted in greater eye fatigue. *Nomad*: Data from highlighted cells in Tables 1 and 2 [16]: R1R1, R2R2, R3R3; V1F1R1, V2F2R2, V3F3R3. *Haploscope*: Data from highlighted cells in Table 1: mR1R1, mR2R2, mR3R3; mR1A1, mR2A2, mR3A3; bR1R1, bR2R2, bR3R3; bR1A1, bR2A2, bR3A3. Table A3 shows the related *F*-tests. Each panel displays 144 data points.

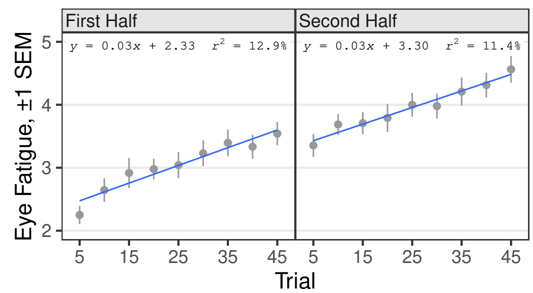

Fig. 6. Across all sessions, as the experiment progressed, there was a linear increase in fatigue. In each session, the level of context switching changed after 45 tasks, separating the session into the first and second halves. Data from all cells in Table 1.

size at the far distance. The Nomad uses a laser-based retinal scanning technology, and all such displays exhibit laser speckle, which reduces image quality and can be particularly problematic for text and graphics with a small visual footprint [8]. In contrast, on the Haploscope text size in terms of visual angle was constant regardless of distance, the image generators did not exhibit laser speckle, and the display resolution was 1920×1080 pixels, compared to 800×600 for the Nomad. Although the maximum tested distance was 4 meters, instead of the 6 meters tested on the Nomad, for most participants 4 meters is still very far from their resting point of accommodation, so this explanation for the different results on the Nomad and Haploscope seems unlikely. Instead, the most likely reason for the increased performance is improved AR image quality. On the Haploscope, increasing reference distance reduced performance, but only under monocular viewing.

Context switching caused greater eye fatigue on the Nomad, and as hypothesized (**H1**), it also caused greater fatigue on the Haploscope. Context switching was more fatiguing at every distance. Given the replication of this effect on two very different display devices, this finding is consistent with the hypothesis that switching cognitive and visual attention between real and AR objects (context switching) causes eye fatigue in all OST AR systems. The effect was stronger on the Nomad than the Haploscope. The effect was also stronger for monocular viewing than for binocular viewing, and in addition, under monocular viewing fatigue increased with increasing distance.

In addition, as shown in Fig. 6 (covering all of the data), as each experimental session progressed, there was a linear increase in reported eye fatigue. As discussed in Section 3.6, participants attended two experimental sessions held on different days, with the viewing condition changing between sessions. In Fig. 6, each experimental session is

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.2022.3150503

broken into two halves, where the level of context switching changed at the half-way point. During this transition, the participant closed their eyes while equipment was moved, which lasted several minutes. After this transition period, eye fatigue declined slightly, but then again began steadily increasing. For both experimental halves, the growth of fatigue occurred at a constant rate of $b = .03$ units per trial. Both slopes significantly differ from 0 (*first half:* $F_{1,214} = 31.8, p < .001^{***}$; *second half:* $F_{1,214} = 27.5, p < .001^{***}$).

4.3 Focal Switching Distance

Larger focal switching distances were expected to reduce performance and increase fatigue (**H2**). The previous section examined context switching when there was no focal distance switching: when *focal switching distance* = 0. Gabbard et al. [16] examined focal distance switching as a binary variable, by comparing cells in which focal distance switching did not occur to cells where it did occur. In the current experiment, this would compare cells with *focal switching distance* = 0 to cells with *focal switching distance* $\neq$ 0. The current analysis instead analyzes focal switching distance as a continuous variable, resulting in more experimental power [41], an analysis of distance switching effects over distance, and an examination of the interaction between context switching and focal distance switching. This section analyzes all of the data in Table 1.

4.3.1 Task Performance

The task performance effects of context switching and focal switching distance are analyzed in Fig. 7. Here, the x axis shows focal switching distance, which takes on the values 0, 1.33, 2, and 3.33 meters (Section 3.4). The columns show the results for *monocular* (left) and *binocular* (right) viewing. The upper row shows performance in terms of *subtasks completed*, while the lower row shows performance in terms of *subtasks correct*.

There was no interaction or main effect of context switching on performance in any panel (Table A4). However, as focal switching distance increased, performance significantly decreased. Under monocular viewing, performance decreased at a rate of $b = .154$ (.163) subtasks per meter completed (correct), while under binocular viewing, performance decreased at a smaller rate of $b = .082$ (.082) subtasks per meter completed (correct). The negative effect of focal switching distance on performance was larger for monocular viewing, $R^2 = 6.9\%$ (5.5%), than for binocular viewing, $R^2 = 2.8\%$ (1.8%).

Table A5 analyzes the effects of viewing and focal switching distance on subtasks completed (correct). Unlike the analysis in the above paragraph, which examines categorical differences within each panel of Fig. 7, this analysis examines categorical differences between panels: between monocular and binocular viewing. There was a significant main effect of viewing, where binocular viewing resulted in $d = .273$ (.358) additional subtasks completed (correct), compared to monocular viewing. These d values are the distances between the lines in Fig. 7. This analysis also finds the main effect of focal switching distance, but the model fits the slopes $b = -.118$ $(-.122)$[3].

4.3.2 Eye Fatigue

The eye fatigue effects of context switching and focal switching distance are analyzed in Fig. 8. Other than the y axis displaying fatigue, the graph structure is the same as Fig. 7. In both monocular and binocular viewing, there was a significant main effect of context switching on fatigue, as well as a significant main effect of focal switching distance (Table A6). Under monocular viewing, context switching increased fatigue by $d = .891$ units, and increasing focal switching distance increased fatigue at the rate of $b = .325$ units per meter. Under binocular viewing, context switching increased fatigue by $d = .507$ units, and increasing focal switching distance increased fatigue at the rate of $b = .196$ units per meter, both smaller amounts. The model explains $R^2 = 22.1\%$ of the variation under monocular viewing, much higher than for binocular viewing, $R^2 = 9.7\%$. Monocular viewing was $d = 1.251$ units more fatiguing than binocular viewing (Table A7).

[3]Note that these are the means of the slopes in panels a, b and c, d in Fig. 7.

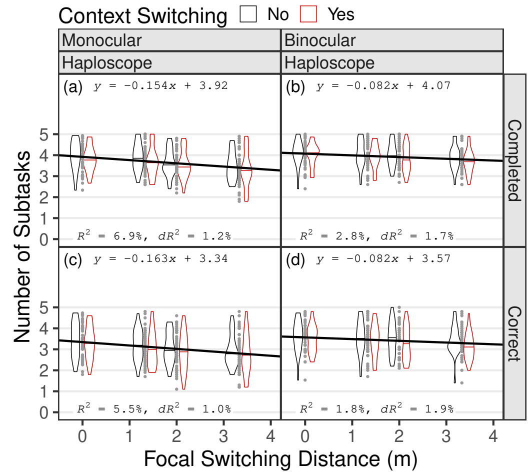

Fig. 7. As focal switching distance increased, performance decreased. Monocular viewing decreased performance. Context switching had no effect. Data from all cells in Table 1. Tables A4, A5 show the related *F*-tests. Each panel displays 192 data points.

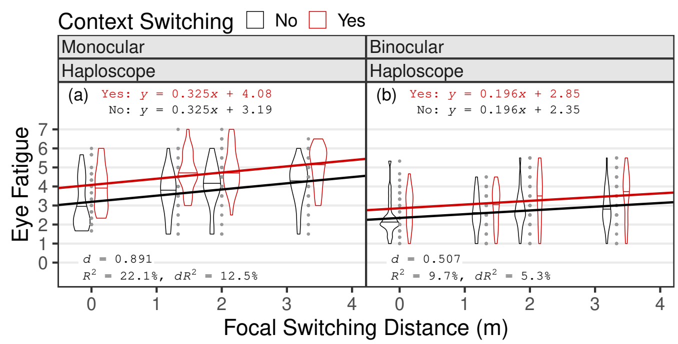

Fig. 8. Context switching increased eye fatigue for both monocular viewing and binocular viewing. Monocular viewing was more fatiguing. Increasing focal switching distance resulted in greater eye fatigue. Data from all cells in Table 1. Tables A6, A7 show the related *F*-tests. Each panel displays 192 data points.

4.3.3 Discussion

In the experiment, participants had to first accommodate to the distance of the left text, and then, if the focal switching distance was greater than 0, change accommodative distance to visually scan the right text. It was therefore hypothesized (**H2**) that increasing focal switching distance would decrease performance and increase eye fatigue. The results support this hypothesis. During focal distance switching, in order to bring information into sharp focus, the eye's ciliary muscles change accommodation, and the eye's vergence muscles change vergence. Therefore, as previously discussed in Section 2.2, continuously shifting eye focus between different focal distances tires these muscles, leading to eye fatigue and reduced performance. In addition, for most people the resting point of accommodation and vergence is about 0.5 meters [49]. In order to accommodate and verge away from the resting point, the eye muscles contract, while when returning to the resting point, the eye muscles relax [19]. Therefore, integrating information closer to the resting point is less exerting. As a result, as the amount of focal switching distance increased, eye fatigue increased and performance decreased.

4.4 Viewing

Viewing has been analyzed in the previous sections, and in each case has been shown to have effects. However, in Section 4.2, while the effects of viewing, context switching, and reference distance were

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.2022.3150503

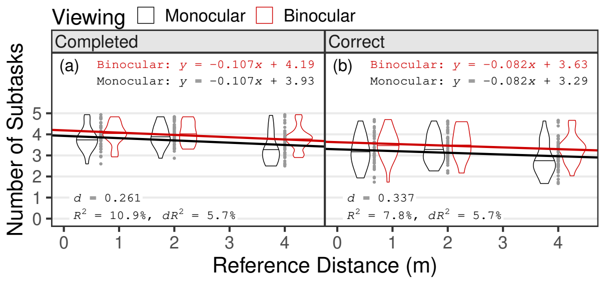

Fig. 9. Task performance was higher under binocular viewing and closer reference distances. Data from all cells in Table 1. Table A10 shows the related *F*-tests. Each panel displays 144 data points.

examined (Figs. 4 and 5), that analysis only covers the 30% of the collected data where focal distance was held constant (focal switching distance $= 0$). In this section, viewing is examined in the context of reference distance, covering all of the data. The structure of the graphs here, Figs. 9 and 10, are the same as Figs. 4 and 5, except that viewing is now analyzed within each panel.

4.4.1 Task Performance

The task performance effects of viewing and reference distance are analyzed in Fig. 9. There was a significant effect of viewing on subtask completion (Fig. 9*a*) and subtask accuracy (Fig. 9*b*). Monocular viewing decreased performance by $d = .261$ $(.337)$ subtasks, describing $dR^2 = 5.7\%$ (5.7%) of the variation. In addition, increasing reference distance resulted in reduced performance, at a rate of $b = .107$ $(.082)$ subtasks per meter. The overall model explains $R^2 = 10.9\%$ (7.8%) of the performance variation.

4.4.2 Eye Fatigue

The eye fatigue effects of viewing and reference distance are analyzed in Fig. 10. There was a significant effect of viewing on fatigue. Monocular viewing increased fatigue by $d = 1.211$ units, a large effect that describes $dR^2 = 33.0\%$ of the fatigue variation. In addition, increasing reference distance resulted in increased fatigue, at a rate of $b = .156$ units per meter. The overall model explains $R^2 = 37.0\%$ of the fatigue variation.

4.4.3 Discussion

When analyzed over all of the data, viewing had the same effects as when it was analyzed over the data where focal distance was held constant (Section 4.2): monocular viewing resulted in lower performance and higher fatigue, and as reference distance increased, performance declined and fatigue increased. However, the effect of viewing was constant, and did not interact with reference distance. Previous work has found that binocular viewing through an HMD provided more accurate accommodation than monocular viewing [28]. As previously discussed in Section 2.1, the primary stimulus that drives the accommodative response is a blur gradient [20]. However, under binocular viewing, stereo disparity additionally drives vergence eye movements, which in turn drive vergence accommodation [27]. So in this experiment, under monocular viewing there was only one accommodative stimuli (blur gradient), while under binocular viewing there were two accommodative stimuli (blur gradient and vergence accommodation). This suggests that under binocular viewing, changing accommodation should be more efficient. In addition, during monocular viewing, participants covered their non-dominant eye with an eye patch, which could have resulted in additional pressure and discomfort. As a result of all of these factors, under monocular viewing performance decreased and fatigue increased.

During the post-experiment informal interview, participants did not complain about image quality at any specific distance, but at or after the midpoint of the experiment, in monocular viewing 7 out of 24 participants complained about general visual fatigue. In contrast, in binocular viewing the number that complained was 2 out of 24. Therefore, participants subjectively reported that monocular viewing was more fatiguing.

4.5 Transient Focal Blur

A performance reduction during focal distance switching, attributed to the transient focal blur effect by Gabbard et al. [16], was expected to replicate under different conditions of context switching and viewing (**H3**). As previously discussed in Section 2.1, changing accommodation from one focal distance to another can be expected to take at least 350 milliseconds, and possibly as long as 425 milliseconds. While accommodation is changing, objects at the new focal distance will be seen with out-of-focus blur. If the task demands performance during this time period, then this *transient focal blur* could cause reduced visual performance.

4.5.1 Letter Undercounts

Gabbard et al. [16] hypothesized that during the transient focal blur time period, participants will be more likely to miss (undercount) target letters in the first line of text. Assuming that participants scan the right text in the standard reading direction of left-to-right, top-to-bottom, the most likely target letters to encounter during this time period would be the letters in the first line of text. Gabbard et al. [16] found that when focal distance switching was required, these letters were significantly more likely to be undercounted. Fig. 11*a* replicates the related graph. Significantly more letters were undercounted per participant when focal distance switching was required, and when a target letter was in the first line of text (Table A8). In addition, there was a significant interaction, where the most letters were undercounted when a target letter was in the first line of text and focal distance switching was required ($p < .001$ for all effects). Gabbard et al. [16] hypothesized that this interaction could be explained by the fact that the task was time-pressured, which caused participants to begin scanning the right text during the transient focal blur period, while their eyes were still accommodating to the new distance.

Hypothesis **H3** is based on the idea that this transient focal blur effect is not specific to AR, but instead is a general property of visual tasks that require integrating information from different displays, possibly located at different focal distances. If correct, this hypothesis predicts that the interaction shown in Fig. 11*a* will replicate under different conditions of context switching and viewing. The current experiment strongly supported this hypothesis: as shown in Figs. 11b,c,d,e, and Table A8, both main effects and their interaction was significant for every combination of context switching and viewing ($p < .001$).

4.5.2 Discussion

Although Gabbard et al. [16] found the transient focal blur effect, their experiment only examined the condition of context switching and semi-binocular viewing of a one-eyed display. This left open the chance that the effect was somehow specific to this set of conditions. However, the replication of the effect, under both conditions of context switching and viewing, suggest that the transient focal blur effect is more general.

When context switching, participants integrated information between a monitor in the real world and an AR display, while when not context switching, both sources of information were monitors in the real world. The transient focal blur effect replicated in both conditions, which is consistent with the hypothesis that the transient focal blur effect is indeed not specific to AR, but is a general property of visual tasks that require integrating information from multiple displays located at different distances.

In addition, the transient blur effect was replicated under conditions of both binocular and monocular viewing. As discussed in the previous section, under binocular viewing, vergence accommodation should make changing accommodation more efficient. There is some evidence for this: when focal distance switching is required, the magnitude of the undercounts is lower with binocular viewing (Fig. 11d,e) than with monocular viewing (Fig. 11b,c). This effect is related to the increased performance and decreased fatigue for binocular viewing discussed in the previous section (4.4). Despite this increased efficiency, the transient focal blur effect was just as statistically strong for binocular viewing ($p < .001$) as it was for monocular viewing ($p < .001$).

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.2022.3150503

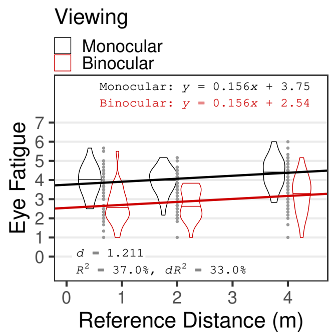

Fig. 10. Monocular viewing was more fatiguing than binocular viewing. Increasing reference distance resulted in greater fatigue. Data from all cells in Table 1. Table A9 shows the related F-tests. The panel displays 144 data points.

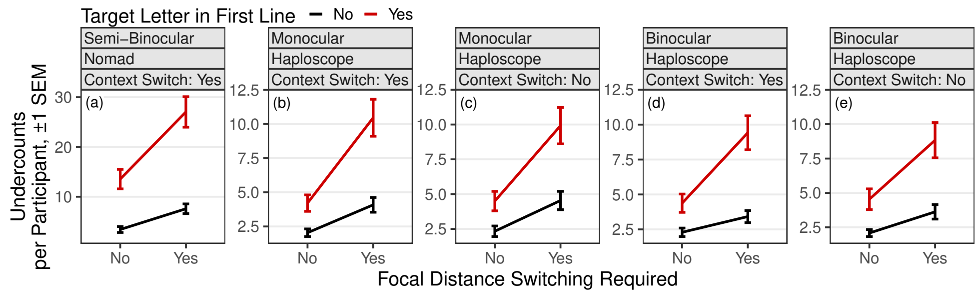

Fig. 11. Evidence for the *transient focal blur effect*: Participants undercounted more letters when a target letter appeared in the first line of the text, and when focal distance switching was required. In addition, these factors interacted; the most letters were undercounted when the target letter was in the first line and participants switched focal distances. In this case, participants tried to read the letter in the first line while that line was still out of focus. (a) The transient focal blur effect found by Gabbard et al. [16]. (b–e) There was a strong transient focal blur effect for every combination of viewing and context switching. *Nomad*: Data from Table 2 (Gabbard et al. [16]): Focal Distance Switching Required: *No* = hatched cells; *Yes* = remaining cells. *Haploscope*: Data from all cell in Table 1: Focal Distance Switching Required: *No* = shaded cells; *Yes* = remaining cells. Table A8 shows the related F-tests.

5 Conclusions and Future Work

This experiment examined the effects of context switching, focal switching distance, binocular and monocular viewing, and transient focal blur. The visual search task required integrating information distributed between real and virtual contexts. The experiment was conducted on a custom-build AR Haploscope, which allowed accurate representation of focal distances and vergence angles. The experiment partially replicated and extended a previous experiment conducted on a Microvision Nomad, a one-eye display viewed semi-binocularly [16]. The primary findings are:

- *Context switching* did not reduce task performance, but did increase eye fatigue.
- As *focal switching distance* increased, performance decreased and eye fatigue increased.
- Compared to *binocular viewing*, *monocular viewing* resulted in reduced performance and increased eye fatigue.
- *Transient focal blur* resulted in reduced task performance under all combinations of context switching and viewing.

Age Effects: This experiment found important effects of accommodation, but used young participants not yet subject to age-related presbyopia [14]. In the current experiment, the mean age of participants was 22.9 years, while in Gabbard et al. [16] the mean age was 22.6 years. It seems reasonable that age would effect performance on the reported task, and may interact in unexpected ways with other experimental conditions. The experiment should be replicated with middle-aged and older participants.

Background Effects: The real world is complex, dynamic, and consists of different colors, objects, shapes, and lighting conditions. In this experiment, white text was presented on a static black background. Future similar studies should examine how different backgrounds interact with context switching and focal distance switching.

Graphical Tasks: Similar to other studies [16, 24, 47], a text-based visual search task was used to examine context and focal distance switching. Advantages of a text-based task include the fact that text is ubiquitous in AR, and many tasks involve comparing or associating text strings. However, graphical elements are also ubiquitous, and so graphical tasks should also be examined.

Tracking Eye Movements: The assumption behind the transient focal blur explanation (Section 4.5.1) assumes that participants scan the right text in the standard reading direction of left-to-right, top-to-bottom. The experiment should be replicated with an eye tracker, which could verify that eye gaze moves in the predicted pattern. This would allow testing the hypothesis that the participant's eye gaze is on the first line of text during the transient focal blur time period. Eye movement data would also enrich the understanding of context switching and focal switching distance effects.

The Transient Focal Blur Effect: Gabbard et al. [16] found that when focal distance switching was required, targets in the first line of text were significantly more likely to be undercounted. This phenomena was attributed to viewing the targets during the transient focal blur time period. The current experiment found additional evidence for the transient focal blur effect, and determined that it exists both when information is distributed between real and AR contexts, and when information is distributed between different real locations. It also exists in conditions of both monocular and binocular viewing. These findings are consistent with the hypothesis that the transient focal blur effect is not specific to AR, but is a general multi-display user interface issue for time-pressured tasks. It is also consistent with the hypothesis that the vergence accommodation available during binocular viewing does not mitigate the effect.

The authors performed a general literature search to see if the transient blur effect has been previously reported. Many studies have explored how different amounts of blur impact the performance of visual search tasks. These show that increasing blur reduces performance [38, 45]. However, the authors could not find previous research that has discussed the effects of transient focal blur, which happens within the short time frame of reaccommodation to a new focal distance. A possible conclusion is that tasks with the properties of what has been examined here—time pressured and requiring integrating information from different displays located at different focal distances—are uncommon in the general field of display design. However, the fixed focal distance of OST AR displays make tasks with these properties likely. Therefore, a conclusion is that these findings on transient focal blur constitute an important addition to knowledge about AR user interface design issues.

Acknowledgments

This material is based upon work supported by the National Science Foundation, under awards IIS-1320909 and IIS-1937565, to J. Edward Swan II. This work was conducted at the Center for Advanced Vehicular Systems, at Mississippi State University.

References

[1] M. S. Arefin. Impact of context switching and focal distance switching on human performance in all augmented reality system. Master's thesis, Mississippi State University, Mississippi State, MS 39762, USA, 2020.

[2] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. Swan II. Impact of AR display context switching and focal distance switching on human performance: Replication on an AR haploscope. In *Abstracts and Workshops Proceedings, IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 2020)*, pp. 571–572. IEEE Computer Society, March 2020. doi: 10.1109/VRW50115.2020.00137

[3] M. D. Barrera Machuca and W. Stuerzlinger. *The Effect of Stereo Display Deficiencies on Virtual Hand Pointing*, p. 1–14. Association for Computing Machinery, New York, NY, USA, 2019.

[4] A. U. Batmaz, M. D. B. Machuca, D. M. Pham, and W. Stuerzlinger. Do head-mounted display stereo deficiencies affect 3d pointing tasks in ar and vr? In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 585–592, 2019. doi: 10.1109/VR.2019.8797975

[5] M. L. Bernard, B. S. Chaparro, M. M. Mills, and C. G. Halcomb. Comparing the effects of text size and format on the readibility of computer-displayed times new roman and arial text. *Int. J. Hum.-Comput. Stud.*, 59(6):823–835, Dec. 2003. doi: 10.1016/S1071-5819(03)00121-6

[6] F. W. Campbell and G. Westheimer. Dynamics of accommodation responses of the human eye. *The Journal of Physiology*, 151(2):285–295, 1960. doi: 10.1113/jphysiol.1960.sp006438

[7] J. H. Carter. *A servoanalysis of the human accommodative mechanism*. PhD thesis, Indiana University, Bloomington, Indiana, USA, 1962.

[8] K. V. Chellappan, E. Erden, and H. Urey. Laser-based displays: a review. *Appl. Opt.*, 49(25):F79–F98, Sep 2010. doi: 10.1364/AO.49.000F79

[9] S. A. Cholewiak, G. D. Love, and M. S. Banks. Creating correct blur and its effect on accommodation. *Journal of Vision*, 18(9):1–1, 09 2018. doi: 10.1167/18.9.1

[10] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken. *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum Associates Publishers, 3 ed., 2003. doi: 10.4324/9780203774441

[11] J. E. Cutting and P. M. Vishton. Chapter 3 - perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein and S. Rogers, eds., *Perception of Space and Motion*, Handbook of Perception and Cognition, pp. 69–117. Academic Press, San Diego, 1995. doi: 10.1016/B978-012240530-3/50005-5

[12] S. Debernardis, M. Fiorentino, M. Gattullo, G. Monno, and A. E. Uva. Text readability in head-worn displays: Color and style optimization in video versus optical see-through devices. *IEEE Transactions on Visualization and Computer Graphics*, 20(1):125–139, Jan 2014. doi: 10.1109/TVCG.2013.86

[13] M. Drouot, N. Le Bigot, J.-L. de Bougrenet, and V. Nourrit. Effect of context and distance switching on visual performances in augmented reality. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 476–477, 2021. doi: 10.1109/VRW52623.2021.00120

[14] A. Duane. Normal values of the accommodation at all ages. *J. of the American Medical Association*, LIX:1010–1013, 1912.

[15] A. Eiberger, P. O. Kristensson, S. Mayr, M. Kranz, and J. Grubert. Effects of depth layer switching between an optical see-through head-mounted display and a body-proximate display. In *Symposium on Spatial User Interaction*, SUI '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3357251.3357588

[16] J. L. Gabbard, D. G. Mehra, and J. E. Swan II. Effects of AR display context switching and focal distance switching on human performance. *IEEE transactions on visualization and computer graphics*, 25:2228 – 2241, 2018. doi: 10.1109/TVCG.2018.2832633

[17] J. L. Gabbard and J. E. Swan II. Usability engineering for augmented reality: Employing user-based studies to inform design. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):513–525, May 2008. doi: 10.1109/TVCG.2008.24

[18] J. L. Gabbard, J. E. Swan II, D. Hix, S.-J. Kim, and G. Fitch. Active text drawing styles for outdoor augmented reality: A user-based study and design implications. In *Technical Papers, Proceedings of IEEE Virtual Reality 2007*, pp. 35–42, March 2007.

[19] H. Gross, F. Blechinger, and B. Achtner. *Human Eye*. John Wiley & Sons, Ltd, 2015. doi: 10.1002/9783527699247.ch1

[20] R. T. Held, E. A. Cooper, J. F. O'Brien, and M. S. Banks. Using blur to affect perceived distance and size. *ACM Transactions on Graphics*, 29(2):19:1–16, Mar. 2010. doi: 10.1145/1731047.1731057

[21] C. Hendrick. Replications, strict replications, and conceptual replications: Are they important? *Journal of Social Behavior and Personality*, 5(4):41–49, 1990. doi: 10.1371/journal.pbio.3000691

[22] G. Heron and B. Winn. Binocular accommodation reaction and response times for normal observers. *Ophthalmic and Physiological Optics*, 9, 1989. doi: 10.1111/j.1475-1313.1989.tb00839.x

[23] 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*, 8(3):1–30, Jan. 2008. doi: 10.1167/8.3.33

[24] A. Huckauf, M. H. Urbina, J. Grubert, I. Böckelmann, F. Doil, L. Schega, J. Tümler, and R. Mecke. Perceptual issues in optical-see-through displays. In *Proceedings of the ACM Symposium on Applied Perception in Graphics and Visualization*, pp. 41–48. ACM, 2010. doi: 10.1145/1836248.1836255

[25] J. H. Iavecchia, H. P. Iavecchia, and I. Stanley N Roscoe. Eye accommodation to head-up virtual images. *Human Factors*, 30(6):703–712, 1988-01-01. doi: 10.1177/001872088803000605

[26] S. Imamov, D. Monzel, and W. S. Lages. Where to display? how interface position affects comfort and task switching time on glanceable interfaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 851–858, 2020. doi: 10.1109/VR46266.2020.00110

[27] D. Kersten and G. E. Legge. Convergence accommodation. *Journal of the Optical Society of America*, 73(3):332–338, Mar 1983. doi: 10.1364/JOSA.73.000332

[28] G.-A. Koulieris, B. Bui, M. S. Banks, and G. Drettakis. Accommodation and comfort in head-mounted displays. *ACM Transactions on Graphics*, 36(4), jul 2017. doi: 10.1145/3072959.3073622

[29] G. Kramida. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1912–1931, July 2016. doi: 10.1109/TVCG.2015.2473855

[30] V. V. Krishnan and L. Stark. A heuristic model for the human vergence eye movement system. *IEEE Transactions on Biomedical Engineering*, 24(1):44–49, 1977. doi: 10.1109/TBME.1977.326207

[31] E. Kruijff, J. E. Swan II, and S. Feiner. Perceptual issues in augmented reality revisited. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2010)*, pp. 3–12. IEEE, Oct. 2010. doi: 10.1109/ISMAR.2010.5643530

[32] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn. Visual discomfort and visual fatigue of stereoscopic displays: a review. *Journal of Imaging Science and Technology*, 53(3):030201–030201–14, Jan. 2009. doi: 10.2352/J.ImagingSci.Technol.2009.53.3.030201

[33] G. D. Love, D. M. Hoffman, P. J. W. Hands, J. Gao, A. K. Kirby, and M. S. Banks. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics Express*, 17(18):15716–15725, 2009. doi: 10.1364/OE.17.015716

[34] W. R. Miles. Ocular dominance in human adults. *Journal of General Psychology*, 3(3):412–430, 1930. doi: 10.1080/00221309.1930.9918218

[35] R. J. Miller, R. G. Pigion, M. F. Wesner, and J. G. Patterson. Accommodation fatigue and dark focus: The effects of accommodation-free visual work as assessed by two psychophysical methods. *Perception & Psychophysics*, 34(6):532–540, 1983. doi: 10.3758/BF03205906

[36] M. Mon-Williams and J. R. Tresilian. Ordinal depth information from accommodation? *Ergonomics*, 43(3):391–404, 2000. doi: 10.1080/001401300184486

[37] M. Mon-Williams and J. Wann. Binocular virtual reality displays: When problems do and don't occur. *Human Factors*, 40:42–49, 03 1998. doi: 10.1518/001872098779480622

[38] A. Musa, A. R. Lane, and A. Ellison. The effects of induced optical blur on visual search performance and training. *Quarterly Journal of Experimental Psychology*, 75(2):277–288, 2022. doi: 10.1177/17470218211050280

[39] C. Neveu, T. Blackmon, and L. Stark. Evaluation of the effects of a head-mounted display on ocular accommodation. *Presence: Teleoperators and Virtual Environments*, 7(3):278–289, 1998. doi: 10.1162/105474698565712

[40] D. A. Owens and K. Wolf-Kelly. Near work, visual fatigue, and variations of oculomotor tonus. *Investigative Ophthalmology & Visual Science*, 28(4):743–749, 1987.

[41] E. J. Pedhazur. *Multiple Regression in Behavioral Research (2nd ed.)*. Holt, Rinehart, & Winston, 1982.

[42] N. Phillips, K. Massey, M. S. Arefin, and J. E. Swan II. Design and calibration of an augmented reality haploscope. In *2018 IEEE International*

Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 75–76. IEEE, Munich, Germany, Oct. 2018. doi: 10.1109/ISMAR-Adjunct.2018.00037

[43] N. Phillips, K. Massey, M. S. Arefin, and J. E. Swan II. Design, assembly, calibration, and measurement of an augmented reality haploscope. In *Proceedings of PERCAR: The Fifth IEEE VR Workshop on Perceptual and Cognitive Issues in AR, IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2019)*, pp. 1770–1774. IEEE, Osaka, Japan, Mar. 2019. doi: 10.1109/VR.2019.8798335

[44] C. Porac and S. Coren. The dominant eye. *Psychological Bulletin*, 83(5):880–897, 1976. doi: 10.1037/0033-2909.83.5.880

[45] A. Schroeger, J. W. Tolentino-Castro, M. Raab, and R. Cañal-Bruland. Effects of visual blur and contrast on spatial and temporal precision in manual interception. *Experimental Brain Research*, 239:3343—-3358, 2021. doi: 10.1007/s00221-021-06184-8

[46] B. D. Schwab, R. E. Hand, and B. N. S. Vlaskamp. Display systems and methods for clipping content to increase viewing comfort. patentus 20190311527A1, Magic Leap Inc., 2019.

[47] B. Schwerdtfeger, R. Reif, G. Klinker, D. Hamacher, L. Schega, and F. Doil. Pick-by-vision: a first stress test. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 115–124. IEEE, 2009. doi: 10.1109/ISMAR.2009.5336484

[48] G. Singh, S. R. Ellis, and J. E. Swan II. The effect of focal distance, age, and brightness on near-field augmented reality depth matching. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–14, 2018. doi: 10.1109/TVCG.2018.2869729

[49] D. R. Tufano. Automotive huds: The overlooked safety issues. *Human Factors*, 39(2):303–311, 1997. PMID: 9302890. doi: 10.1518/001872097778543840

[50] K. Ukai and P. A. Howarth. Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. *Displays*, 29(2):106–116, 2008. doi: 10.1016/j.displa.2007.09.004

[51] E. Viirre, R. Johnston, H. Pryor, S. Nagata, and T. A. Furness. Laser safety analysis of a retinal scanning display system. *Journal Of Laser Applications*, 9:253–260, 1997. doi: 10.2351/1.4745467

[52] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '00, p. 110–119. Association for Computing Machinery, New York, NY, USA, 2000. doi: 10.1145/345513.345271

[53] R. W. West and D. A. Goss. *Introduction to the Optics of the Eye*. Butterworth-Heinemann, 2001.

[54] M. Yeh, C. Swider, Y. J. Jo, and C. Donovan. Human factors considerations in the design and evaluation of flight deck displays and controls, version 2.0. Technical Report DOT-VNTSC-FAA-17-02, Federal Aviation Administration, Human Factors Division, 2016.