Understanding, Addressing, and Analysing Digital Eye Strain in Virtual Reality Head-Mounted Displays

TERESA HIRZLE, FABIAN FISCHBACH, JULIAN KARLBAUER, and PASCAL JANSEN,
Institute of Media Informatics, Ulm University
JAN GUGENHEIMER, Institut Polytechnique de Paris, Télécom Paris - LTCI
ENRICO RUKZIO, Institute of Media Informatics, Ulm University
ANDREAS BULLING, Institute for Visualization and Interactive Systems, University of Stuttgart

Digital eye strain (DES), caused by prolonged exposure to digital screens, stresses the visual system and negatively affects users' well-being and productivity. While DES is well-studied in computer displays, its impact on users of virtual reality (VR) head-mounted displays (HMDs) is largely unexplored—despite that some of their key properties (e.g., the vergence-accommodation conflict) make VR-HMDs particularly prone. This work provides the first comprehensive investigation into DES in VR HMDs. We present results from a survey with 68 experienced users to *understand* DES symptoms in VR-HMDs. To help *address* DES, we investigate eye exercises resulting from survey answers and blue light filtering in three user studies (N = 71). Results demonstrate that eye exercises, but not blue light filtering, can effectively reduce DES. We conclude with an extensive *analysis* of the user studies and condense our findings in 10 key challenges that guide future work in this emerging research area.

CCS Concepts: • **Human-centered computing** → *Empirical studies in HCI*; **Virtual reality**; • **Applied computing** → **Consumer health;**

Additional Key Words and Phrases: Virtual reality, digital eye strain, well-being

ACM Reference format:
Teresa Hirzle, Fabian Fischbach, Julian Karlbauer, Pascal Jansen, Jan Gugenheimer, Enrico Rukzio, and Andreas Bulling. 2022. Understanding, Addressing, and Analysing Digital Eye Strain in Virtual Reality Head-Mounted Displays. *ACM Trans. Comput.-Hum. Interact.* 29, 4, Article 33 (March 2022), 80 pages.
https://doi.org/10.1145/3492802

This work has been supported by the DFG project "Empirical Assessment of Presence and Immersion in Augmented and Virtual Realities" (RU 1605/4-1) and by the Landesgraduiertenförderung (LGFG) Scholarship for PhD students. A. Bulling was funded by the European Research Council (ERC; grant agreement 801708).
Authors' addresses: T. Hirzle, F. Fischbach, J. Karlbauer, P. Jansen, and E. Rukzio, Institute of Media Informatics, Ulm University, James-Franck-Ring 1, Ulm, 89081, Germany; emails: {teresa.hirzle, fabian.fischbach, julian.karlbauer, pascal.jansen, enrico.rukzio}@uni-ulm.de; J. Gugenheimer, Institut Polytechnique de Paris, Télécom Paris - LTCI, 19 Place Marguerite Perey; email: jan.gugenheimer@telecom-paris.fr; A. Bulling, Institute for Visualization and Interactive Systems, University of Stuttgart, Pfaffenwaldring 5a; email: andreas.bulling@vis.uni-stuttgart.de.

1073-0516/2022/03-ART33 $15.00
https://doi.org/10.1145/3492802

1 INTRODUCTION

A large body of work has shown that prolonged exposure to digital screens causes vision and eye problems [12, 24, 79], collectively referred to as ***Digital Eye Strain (DES)*** or ***Computer Vision Syndrome (CVS)*** [80]. Symptoms of DES include, but are not limited to, dry eyes, headache, double, or blurred vision [86], and can lead to negative impacts on a person's general well-being, quality of life [64], and productivity [79]. The high demand for near-vision responses when looking at conventional displays may further lead to accommodative fatigue [80] and can cause changes in vergence and accommodation[1] responses [12]. Whereas in the past only office workers were affected, DES has become a pervasive problem in today's digital society with up to 90% of computer users suffering from the symptoms on a regular basis [79].

DES symptoms are expected to be even more severe in **virtual reality head-mounted displays** (**VR-HMDs**) due to technical characteristics of these devices. This includes problems with depth presentation caused by the **vergence-accommodation conflict** (**VAC**) [42, 59, 94] but also problems caused by the stereoscopic displays, in particular, binocular disparity or stereoscopic distortions [59]. In addition, VR-HMDs cover a wider part of the field of view than conventional displays on which content is usually presented in the foveal area. Presentation of content also in the periphery might lead to larger saccades, further increasing eye strain [38].

Especially when considering the broader distribution of VR-HMDs to consumers in recent years, it is expected that DES symptoms will become even more prevalent and severe. However, while DES is well-studied with conventional displays, such as laptops or smartphones [80], VR-HMDs have received only little attention so far. Especially in comparison to the well-studied problem of simulator sickness in VR, the issue of DES is insufficiently studied, despite that recent research suggests that users might, in fact, at least be equally affected by both problems [39]. Furthermore, existing approaches to address DES in VR-HMDs have primarily focused on the VAC as the main cause of symptoms [38].

This work presents the first fundamental and comprehensive investigation into DES in VR-HMDs by covering three essential parts (see Figure 1). We first present an online survey with 68 current VR- HMD users to better *understand* how they are affected by DES and what strategies they use to cope with it. The survey confirmed that DES is, in fact, a widespread and unsolved problem for VR-HMD users. A total 46% of the users reported experiencing symptoms (e.g., headache, dry eyes, and increased sensitivity to light) at least once per hour of usage, and 50% are concerned about VR-HMDs harming their eyes. To mitigate or reduce DES, some users apply coping strategies, such as closing the eyes or blinking quickly. However, most of them simply interrupt device use. These results indicate that device-integrated solutions, which can be applied during device use and do not require users to interrupt their experience, are missing. Inspired by these results, we investigated two techniques to *address* DES in VR-HMDs that can be applied during device use. For the first technique, we implemented two versions of the blue light filter method. Blue light filtering is broadly spread among users of conventional digital devices to reduce DES symptoms and some VR-HMD manufacturers already offer it as "night mode" in their devices, e.g., for the Oculus Quest 2.[2] However, an empirical investigation of the effect of blue light filters in VR-HMDs is currently missing. For the second technique, we designed a set of eye exercises grounded in the insights of the survey. The exercises consist of short (30 sec) visual tasks and are intended to mitigate the major symptoms of DES. We investigated the effectiveness of these approaches in three

[1]Vergence and accommodation refer to two mechanisms of the visual system to perceive and process depth information. Vergence refers to the simultaneous inward rotation of the eyes to fuse the images that both eyes perceive to one percept, while accommodation refers to the bending of the eye lens to bend the entering light onto the fovea [27].

[2]© Facebook Technologies, LLC.: https://www.oculus.com/quest-2/, last retrieved: April 27, 2021.

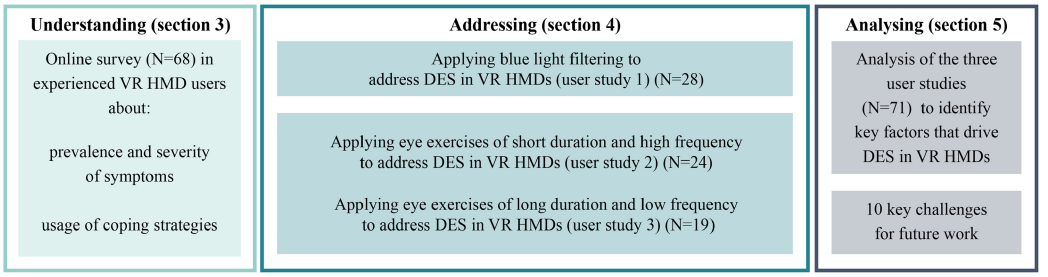


Fig. 1. This work investigates DES in virtual reality head-mounted displays from three complementary perspectives. In the first part *understanding* we report results of a survey about the prevalence and severity of symptoms in experienced VR-HMD users. We also asked them about coping strategies that they use to alleviate the problem. In the second part, we investigate two solutions to *address* the problem of DES in VR-HMDs in three user studies. The third part, *analyzing*, includes an analysis of the three user studies, revealing that the factors of sex and susceptibility drive DES in VR-HMDs. We conclude with 10 key challenges that guide future research on DES in VR-HMDs.

user studies. In the first user study (N = 28), we compared the potential effects of two versions of blue light filtering with a control condition. In the second and the third user study (N = 24 and N = 19), we investigated the frequency and duration of eye exercises. In all three user studies, we found that participants experienced DES and that symptoms increased significantly in a usage time of 25 minutes. We found that the set of eye exercises significantly reduced DES symptoms directly after their application and had a significant extended effect even after a second straining VR task compared to a control condition. However, we did not find positive effects of the blue light filter on DES symptoms. We conclude with a comprehensive *analysis* of all three user studies (N = 71). This analysis revealed that women were slightly more severely affected by DES than men across all three user studies. Furthermore, we identified that participants could be grouped into two sensitivity groups using a clustering analysis, with one group experiencing symptoms significantly more severe than the other. Based on these analyzes, we present 10 key challenges that guide future work in investigating and alleviating DES in VR-HMDs. In these key challenges, we highlight the need to conduct further investigations into DES to get a more nuanced and holistic understanding of the different types of discomfort than can occur as negative side effects of VR-HMD use. Furthermore, we argue that DES measurement should become part of the landscape of measurements applied to evaluate VR experiences and devices. Lastly, we point out how long-term challenges or the relation to other VR usability metrics could impact future research.

2 RELATED WORK

This section gives a general overview of the causes and symptoms of DES, both with conventional digital displays and specific to HMDs. Furthermore, we discuss objective and subjective measures of DES. Lastly, we show how our research is related to previous work on approaches to alleviate DES.

Terminology. In the literature, different terms are used interchangeably to refer to vision and eye problems resulting from prolonged exposure to digital screens [38]. The formal diagnostic term for *eye strain* is *asthenopia* [86] and refers to "complaints related to refractive error or ocular muscle imbalance" [68]. The effects of eye strain can be divided into the two components *visual fatigue* and *visual discomfort* [59]. *Visual fatigue* refers to implications of the performance of visual functions

and is measured technically, e.g., with optometric instruments. On the other hand, *visual discomfort* refers to the subjective component that is assessed with subjective user ratings [59]. *DES* is used synonymously with the term *CVS* and refers to the type of asthenopia that specifically stems from digital device usage [80].

2.1 Causes

2.1.1 General Causes of DES. Since the first introduction of video display terminals and later computers at the workplace, many studies have investigated potential negative effects on users' eye health (e.g., [26, 31, 47, 90]). Besides the number one cause exposure time, several additional causes were suggested, including factors that stem from problems of the visual system (e.g., astigmatism), properties of displays, or environmental settings [12, 24]. Coles-Brennan et al. proposed an approach to structure causes into five categories: vision-related, oculomotor-related, ocular surface-related, environmental factor-related, and device-related [24]. Vision-related causes stem from vision problems, such as refractive error or presbyopia [87]. Oculomotor-related causes are caused by disturbances in oculomotor responses, e.g., fixation disparity [25]. Ocular surface-related causes result from irritations of the ocular surface of the eye. Besides environmental reasons or contact lenses [56], this is assumed to be mainly caused by dry eye, e.g., a reduced blink rate [80, 90]. Environmental factors are related to the environment where the device is being used and include lighting or humidity [12]. Lastly, there is a large set of device-related causes, such as close viewing distances that, when held for a prolonged time period, can cause a high demand of vergence and accommodation responses and, as such, tensions in the eye muscles [79, 94]. Other display-based factors are glare [12], screen brightness [53], or color [101]. Besides this set of passive causes, particular interaction techniques or applications that require unnatural oculomotor responses, such as prolonged fixation duration [93] or many long saccades [6], promote the occurrence of eye strain. In practice, this may happen if a user interface requires users to keep their gaze on an element for an extended amount of time (e.g., dwell-time interaction [18, 66]), if the interface requires a lot of eye movements for the user to find and select the right user interface element [9, 17], or when the user interface requires multiple gaze commands [74]. While researchers attempted to isolate causes and measure their influence on symptoms experimentally, currently, no transparent model exists that links symptoms and causes of DES [79].

2.1.2 Causes Specific to VR-HMDs and Stereoscopic Displays. Besides general causes of DES, stereoscopic displays and HMDs incorporate several properties that promote eye strain further. The most prominent cause for DES in HMDs is the VAC, i.e., the decoupling of vergence and accommodation responses due to conflicts in depth representation [40, 42, 94]. In a recent study, Vienne et al. found that vergence responses were slower immediately after the exposure to conflicting ocular depth cues [94]. Therefore, the authors suggest using these vergence dynamics as objective indicators of DES in stereoscopic viewing conditions. Some attempts were made to measure strain that stems particularly from this conflict, e.g., by relating **Electroencephalography** (**EEG**) signals to **Simulator Sickness Questionnaire** (**SSQ**) scores [62]. But given that there are yet more factors presumed that make HMDs particularly prone to DES, it is difficult to isolate effects that stem solely from the VAC. For instance, an incorrect setting of interpupillary distance can cause blurry and double vision [102]. Furthermore, influences like blur, glare, or refresh rate might become more severe when applied over a larger field of view. All these indicators precisely point toward current HMD technology, causing an increased occurrence of DES. However, to the best of our knowledge, an analysis of the prevalence and severity of symptoms during natural usage behavior of VR-HMD users in their homes is currently missing.

2.2 Symptoms

2.2.1 General Symptoms of DES. DES is polysymptomatic and affects the visual system with several problems that can have extended effects on a person's general well-being and quality of life [64]. The set of symptoms is extensive and includes eye-related as well as extraocular symptoms. Extraocluar symptoms include neck pain, shoulder pain, headache, or backache [61]. Eye-related symptoms can be grouped into being perceived inside the eye or externally [86]. Sheedy et al. investigated the effects of several inducing conditions of DES (e.g., close viewing distance or dry eyes) to symptoms [85]. They found that internal symptoms (strain, ache, and headache) are related to inducing conditions that stem from visual functions (accommodative or convergence stress or refractive error). On the other hand, external symptoms (burning, irritation, and dryness) are caused by an irritation of the corneal surface of the eyes. Further studies confirmed this distinction between internal and external eye-related symptoms (e.g., Zeri and Livi [103]).

2.2.2 Symptoms Specific to HMDs and Stereoscopic Displays. Since their introduction, researchers investigated the temporary and long-term effects of VR-HMDs on the visual system that would occur due to mismatches between the real and the simulated world of an HMD [72]. In an early study, Peli et al. did not find differences in visual functions but in subjective ratings after a 30-min exposure to a stereoscopic VR-HMD compared to a desktop computer [71]. Participants rated the VR-HMD exposure less comfortable than the exposure to the desktop display. Other studies that assessed subjective DES in a comparison of HMDs to certain display types (e.g., tablets [99], TV monitors [58] or desktop, and projection displays [84]) support these findings that HMDs cause a considerable amount of more discomfort. However, these studies were conducted some years ago when HMDs were heavier and bulkier than today. Besides the form factor, the literature suggests that the stereoscopic presentation of content in HMDs is a solid contributor to subjective DES [58, 71]. Zeri and Livi [103] investigated the type of symptoms in stereoscopic displays and found a similar structure of internal and external symptom factors as indicated by Sheedy et al. [86]. While Sheedy et al. [86] found the sensation of external symptoms was higher than internal ones, Zeri and Livi [103] report the opposite. These results could be explained by the VAC present in stereoscopic displays. The VAC might increase accommodative stress and, with it, internal symptoms, which are directly related to it.

2.2.3 Difference in Simulator Sickness and DES Symptoms. Currently, simulator sickness is the dominant discomfort type investigated with VR-HMDs [39]. The difference between the two concepts, simulator sickness and DES, is not entirely clear because simulator sickness is dominantly measured with the SSQ [50], which includes an oculomotor sub scale covering symptoms that have also been attributed to DES (e.g., eye strain or blurred vision). Therefore, eye strain has often been considered a sub-category of simulator sickness [38]. Previous works revised the symptoms that occur in VR. For example, Ames et al. developed a questionnaire explicitly addressing symptoms that stem from VR viewing, covering 11 ocular and 12 non-ocular symptoms [4]. Kim et al. proposed the VR sickness questionnaire, an adapted version of the SSQ specifically for VR use [51]. Both of these works consider oculomotor symptoms an important part of VR symptomatology but consider them a sub-category of simulator sickness, and as such, do not provide a clear distinction between the two constructs. In contrast, in a recent study, Hirzle et al. provide a distinction of the two concepts. They divide the discomfort that is experienced with VR-HMDs into three factors: simulator sickness, DES, and ergonomic symptoms [39]. The simulator sickness factor includes symptoms relating to a feeling of nausea, such as dizziness, vertigo, or fullness of head. The DES factor includes eye and vision-related symptoms, such as dry eyes, eye ache, or irritation.

Interestingly, they did not find an overlap between the two factors suggesting that the concepts are indeed different.

2.3 Measures

Due to the omnipresence of DES in today's life, Rosenfield argued that DES assessment should become an integral part of today's standard eye examinations [80]. Besides external and internal symptoms, DES can be decomposed into an objective and a subjective part, as pointed out by Lambooji et al. [59]. The authors name *visual fatigue* as a term that refers to the objective component measured technically by using an eye tracker and *visual discomfort* that describes the subjective element of DES, assessed by subjective user ratings.

2.3.1 Objective Measures. While optometric instruments provide high quality and detail in the measurement of eye properties and visual functions, such as accommodative or vergence responses, it is difficult to use them without the expertise of an optometrist. Therefore, Wang et al. proposed to use eye movements analysis, such as blink metrics, to determine the eye fatigue level instead [97]. At this, blink metrics were by several other studies considered an indicator for the DES in general [30, 73, 89]. However, one has to consider that objective eye measures, such as blink rate or pupil size, are difficult to use in an applied setting, as influence factors such as cognitive or affective load might have a huge impact on these metrics [70, 92].

2.3.2 Subjective Measures. In contrast to objective measures that require optometric expertise or special devices, subjective measures are easier to integrate into study designs for evaluating users' symptoms at home. Proposed questionnaires and measurement scales vary in detail, both of scale and items. In 1985, Howarth and Istance report findings of a two-year study to investigate the prevalence and severity of visual problems that occur with the use of visual display units [44]. For visual discomfort, they considered seven symptoms, rated on a 5-point scale from *no discomfort* to *very bad discomfort*. The symptoms were "tiredness of the eyes", "soreness or aching of the eyes", "soreness or irritation of the eyelids", "watering of the eye", "dryness of the eyes", "a sensation of hot or 'burning' eyes", and "a feeling of 'sand in the eyes"'. Seguí et al. proposed the **Computer Vision Syndrome Questionnaire** (**CVS-Q**) as a consistent measurement tool of the CVS at the workplace [28]. In contrast to other works, the authors propose to assess symptoms on two scales, a frequency (0: never, 1: occasionally, and 2: often or always) and an intensity scale (1: moderate, 2: intense), as they aim at detecting, whether users are affected by the CVS during a time period or not at a single point in time. Sheedy et al. proposed to use visual analog scales to assess the values of nine symptoms (external and internal ones) of DES [86]. One of the most used questionnaires to assess DES and, in general, discomfort in VR-HMDs [39], is the SSQ [50]. Although it includes four oculomotor symptoms (headache, eye strain, difficulty focusing, and blurred vision), it is rather unspecified compared to the purposeful questionnaires mentioned. In a recent study, Hirzle et al. used 21 symptoms to assess DES compared to simulator sickness and ergonomic symptoms in VR-HMDs [39]. Using a factor analysis, they reduce this set of symptoms to 11 symptoms that carry the most critical information on DES. To conclude, in the literature, there seems to be no established questionnaire to evaluate DES specifically in virtual or augmented reality HMDs. In our user studies, we measure DES symptoms subjectively by using the nine symptoms that were proposed by Sheedy et al. [86]. We extend this set of symptoms by two symptoms ("difficulty focusing" and "sensitivity to bright light" from CVS-Q [28] and SSQ [50]).

2.4 Strategies to Alleviate DES

2.4.1 Blue Light Filtering. Due to the polygenic and polysymptomatic nature of DES [38, 79], solutions to reduce specific DES symptoms are challenging to design. A broadly propagated approach

to creating a more comfortable viewing experience with digital devices is blue light filtering. While it has been shown that the intensity of blue light emitted from digital devices is too low to cause damaging effects to the eyes [87], it has been suggested that blue light glasses or filters might help reduce DES [22, 46]. Cheng et al. report a reduction of subjective eye strain symptoms in subjects that suffer from dry eye when wearing blue light filtering glasses compared to a control group [22]. However, a control condition is missing, and as such, results have to be considered with reservations. In Ide et al.'s work, the potential positive effects of blue light filtering lenses are attributed to reduced **critical flicker frequency** (**CFF**) values, but not to significant differences in subjective DES [46]. In addition, recent studies that investigated the effects of blue light filtering contact lenses [81] or screen filters [69, 76] did not find significant effects on subjectively reported eye strain symptoms in a reading task. These results suggest that blue light filtering in digital devices does not significantly impact subjectively perceived eye strain symptoms (at least during exposure of 30 minutes). On the other hand, in recent years, HMD manufacturers have advertised these filters by offering a "night mode" on their devices, e.g., Oculus Quest 2[2]. However, currently, no experimental investigation of blue light filtering in VR-HMDs exists. We investigated the effect of blue light filtering on subjective DES in our first user study.

2.4.2 Technical Approaches. Several technical devices have been proposed to treat dry eye syndrome [2] and strengthen the eye muscles to relieve eye strain [45, 91, 95]. For instance, Bonham and Rallison proposed a holographic system that aims at exercising the ciliary muscles that control accommodation to reduce DES by varying focal distances [13]. However, these systems are not empirically evaluated, and, therefore, their effectiveness is unknown. Furthermore, a variety of technical solutions to alleviate DES were proposed for the use of screen-based digital devices (e.g., when interacting with a computer or mobile screens [41]), often as part of correcting the ergonomic posture [19, 57]. For example, one approach is to apply external stimuli to remind users to keep a healthy distance to the screen [65] or to increase the blink rate [29]. In contrast to these works, we are interested in evaluating the effectiveness of device-integrated solutions that can directly be integrated into VR-HMDs and do not require hardware or software modifications.

2.4.3 Eye Exercises. Another type of alleviation strategy is known in the field of vision therapy. Here, eye exercises are applied to help with eye conditions like convergence insufficiency, heterophoria, or strabismus [43], but also general well-being [75]. In a recent study, Thai et al. suggest the use of eye exercises to reduce symptoms of DES during VR-HMD usage [88]. The authors used two exercises from clinical context and compared their application in VR to the application outside of VR with a control condition. The authors did not find statistically significant differences between the conditions, although descriptively, symptoms were slightly more severe in the control condition.

Other types of visual exercises were presented to improve visual acuity for patients with amblyopia by leveraging the possibility in VR to display different images to the eyes [10, 11, 36]. These systems utilize the unique properties of HMDs to use as vision training devices for particular user groups. In contrast, our work studies risk to eye health relevant to all users of VR-HMDs, and are not specifically tailored toward a pathological issue or a specific user group.

3 UNDERSTANDING DES IN VR HMD USERS

Previous works typically investigated DES in laboratory studies, where DES symptoms were proactively induced to evaluate causes and symptoms. Furthermore, the studies that exist and investigate natural usage behavior were conducted with computer or tablet displays. However, how VR-HMD users are affected during natural usage is currently unknown. Besides, it is unclear how

users cope with DES, as most of the discussed alleviation techniques require hardware or software modifications.

To address these limitations, we consulted users who frequently use VR-HMDs about their DES symptoms and the coping strategies they developed during natural user behavior. We, therefore, designed an online survey that covered the two essential parts *prevalence and severity of symptoms*, and *coping strategies*, summarized by the following research questions:

RQ1: What is the prevalence and severity of DES symptoms in everyday users of VR HMDs?
RQ2: Do users of current VR HMDs apply strategies to mitigate or prevent symptoms?

The survey consisted of 53 questions addressing frequency and duration of device usage, the occurrence of symptoms of DES, and participants' strategies to cope with symptoms. The complete questionnaire is presented in appendix A.

3.1 Questionnaire

Devices and Usage. This first section of the questionnaire included seven questions about which devices the respondents used, how often they used them, how long they had been using their device, and how long they usually used their device in one session. We further asked how often they usually interrupted their VR-HMD session and why they interrupted it. Lastly, we asked when participants had used their headset the last time and whether they used one that allowed them to adjust interpupillary distance.

Eye-Specific Demographics. In the second part, we asked participants three questions on whether they had any vision problems or impairments and whether they used vision correction in their daily lives and VR.

General DES Experience. In the third section, we asked participants eight questions about their general experience with DES. We asked them how often and how intensely they usually experienced DES symptoms *during* and *after* a VR session and after which time period symptoms usually subsided. We also asked them to compare their DES experience in VR with that of a desktop monitor. Lastly, we asked respondents to report a specific scenario in which they experienced DES. In the questionnaire, we used the term "visual discomfort" to refer to all DES symptoms that participants experienced.

Computer Vision Syndrome-Questionnaire. In this section, we assessed specific symptoms of DES. We were interested in capturing participants' experience with DES not only for one specific experience but in general. Therefore, we used the (CVS-Q), which has the purpose of assessing symptoms over a specific time period and indicating whether computer workers are affected by the CVS [28]. The CVS-Q incorporates a frequency and an intensity scale to determine the severity of symptoms. The frequency scale is anchored around the time period of *one week* as a reference. As we were looking for users who use their device frequently, but not eight hours a day, we defined our reference time period as *one hour* of usage. Therefore, we adapted the frequency scale of the CVS-Q by defining "occasionally" as "sporadic episodes or *once per hour of usage*", instead of "sporadic episodes or *once a week*". Similarly, we changed "often or always" from "2 or 3 times *a week* or almost *every day*" to "2 or 3 times *per hour of usage* or almost *every time*". We did not change the intensity scale that covers two intensities, "moderate" and "intense" (and "N/A" if "never" is chosen on the frequency scale).

Coping Strategies. In this section, we asked participants three questions about their knowledge of coping strategies for DES, and whether they applied them or any other methods to cope with DES symptoms.

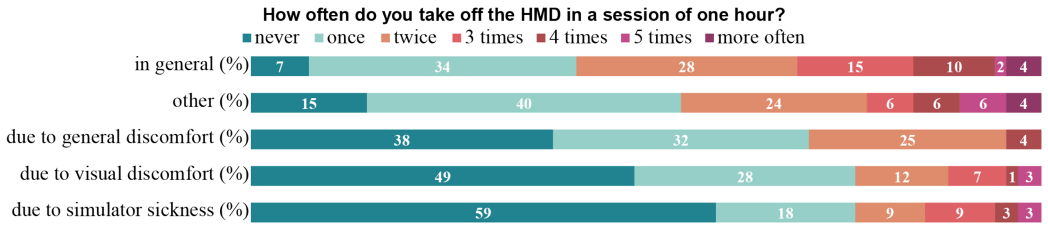


Fig. 2. An overview of how often the online survey respondents indicated to interrupt their VR-HMD sessions of one hour. We found that 93% of the respondents interrupt device use at least once in general, 62% interrupt it at least once due to general discomfort, 51% due to symptoms of DES, 41% due to symptoms of simulator sickness, and 85% due to other reasons.

Usage of Coping Strategies. In the last section, we asked participants three questions on whether and under which circumstances they could imagine using coping strategies to reduce DES symptoms.

3.2 Participants

We conducted the study as an online survey, using the recruiting platform Prolific[3] to recruit participants. We screened participants with a registration study and invited only participants who owned and used a VR-HMD regularly (at least once a week) to participate in the main study. Of the 197 participants who participated in the registration study, 71 used a VR headset regularly and therefore met the criteria for the main study. Finally, 68 participants completed the main study, with a mean time of 13.2 minutes ($SD = 6.7$). On average, participants were 27 years old ($SD = 7.6$). Out of them 29 identified as female and 39 as male. Participants received a reward of £0.63 for the registration and £3.35 for the main study.

3.3 Results

3.3.1 Devices and Usage. A total 91% of the participants stated to own a VR headset, with the most commonly used device being Sony PlayStation VR (38.2%), followed by the Oculus Rift (19.1%), the Samsung GearVR (16.2%), and the HTC Vive (10.3%). Total 6% of the participants had only been actively using their headset for less than a month. Total 43% had actively been using their headset between one and six months and 29% between six and 12 months. Total 22% had actively been using it for two or more than two years. Half of the participants (50%) usually used their device for 30–60 minutes per session. Total 34% of the participants usually used it between 1–2 hours, 6% between 2–3 hours, and only 1% used it for more than 3 hours. Total 9% usually used it for less than 30 minutes in one session. A large part (39.7%) did not know whether the headset they used allowed for the adjustment of the interpupillary distance (IPD). Less than 1% of the participants stated that they adjust IPD most of the time when they use their device, 13% responded to adjust it when they notice some symptoms of eye strain, and 18% stated to have adjusted IPD once at the beginning of using the device. More than half of the respondents (51%) stated that they interrupted their VR session at least once an hour due to symptoms of visual discomfort, with 11% taking the headset off between three and five times (see Figure 2). The number of interruptions for visual discomfort was even higher than for simulator sickness, for which 41% of the participants stated to take off the headset more than once in an hour, with 15% taking it off between three and five times.

[3]Prolific: https://prolific.co/, last retrieved: March 24, 2021.

3.3.2 Eye-Specific Demographics. In total, 51% of the participants stated to have vision problems, i.e., nearsightedness (46%), farsightedness (4%), and astigmatism (12%). Of these 51%, 92% indicated using prescription glasses (49%), contact lenses (9%), or both (34%) to correct their vision problems. Of the respondents who used a form of vision correction, 28% indicated using prescription glasses, and 25% indicated using contact lenses when using VR. In summary, 47% of the participants who indicated a vision problem and usually used vision aids to correct their vision did not use a vision aid in VR (15 participants in total).

3.3.3 General DES Experience. We first asked participants about their general experience with visual discomfort *during* and *after* a VR session. We used the frequency (never, occasionally, often, or always) and intensity (none, moderate, and intense) scales of the CVS-Q to assess this information. Total 37% of the respondents stated that they occasionally experienced visual discomfort during a VR session when asked whether they experienced it in any form. A fourth of this 37% were participants who usually wore vision aids to correct their vision but did not use them in VR. Total 10% of all participants indicated that they often or always experienced visual discomfort, with 6% being participants who did not wear their vision aids in VR. When asked about visual discomfort that occurred after a VR session, 34% stated to experience moderate, visual discomfort, occasionally, and 1% always. When asked after what period of time visual discomfort usually occurred, on average, respondents stated that symptoms started to appear after 32 minutes of using the device. When asked how long it usually takes for the symptoms to subside, 14% named an hour or more after usage, 7% within 30 minutes, 23% within 15 minutes, and 33% within five minutes after use. We further asked participants to compare the frequency of symptoms during VR usage with those that arose while using a desktop screen. The majority (54%) stated that symptoms occurred more often when using VR, 24% when using a desktop screen, and 22% equally often.

We asked participants about one specific scenario that they remembered in which visual discomfort occurred and how they dealt with it. We received 38 answers that two of the authors analyzed using thematic analysis following Braun and Clarke's six-phase process [15]. The goal of thematic analysis is to reveal common themes in the analyzed data. We used a data-driven approach to extract themes in respondents' answers looking for semantic themes that were already present in the data, not trying to find latent themes. As the scenario we asked for was limited per definition, both authors found a set of similar themes. In the following, we summarize the data under these themes. The result of this analysis were three main themes *symptoms*, *causes*, and *strategies*. The most frequent symptoms we found were *discomfort* (8 times), *nausea* (7 times), and *headache* (6 times). Other less frequent symptoms were *strain* (3), *eye pain* (3), and *difficulty focusing* (3). The most common causes were *long exposure* (11 times) and *rapid movement* (7 times). The most common avoidance strategy was to stop using the device (17 times), followed by laying down (3 times, not a sub set of stopping device usage), and closing the eyes for a while (3 times).

3.3.4 Computer Vision Syndrome-Questionnaire (CVS-Q). We asked participants to fill out the adapted CVS-Q to gain information about their normal state of eye strain. The authors of CVS-Q indicate that a score above six signals that a person probably suffers from CVS [28]. We excluded four participants due to contradictory answers (i.e., they reported that a symptom never occurred but marked it as intense on the intensity scale) for calculating the mean score. Total 53% of the participants reached a CVS score of 6 or higher, and the mean score for the CVS-Q was 7.02($SD = 6.1$) on a scale of $0 - 64$. We calculated the single symptom scores by multiplying the intensity and the frequency values for each symptom and participant. We then averaged the single symptom scores for each symptom and found the highest single symptom score for *headache* ($M = 0.97, SD = 1.10$),

Table 1. Survey Respondents' Ratings for the Single Symptoms of the CVS-Q

	Headache	Excessive blinking	Dryness	Blurred vision	Increased sens. to light	Eye redness	Eye pain	Itching	Heavy eyelids	Tearing	Diffic. focusing for near vision	Feeling that sigh is worsening	Double vision	Burning	Feeling foreign body	Col. halos
never (%)	38	55	52	58	58	63	64	64	66	69	67	84	83	83	89	89
occ. mod. (%)	44	28	38	30	31	31	31	33	31	25	31	9	11	16	9	9
occ. intense (%)	8	6	2	3	6	2	2	0	0	3	0	0	5	0	0	0
always mod. (%)	3	6	5	5	3	3	2	3	3	3	2	2	0	2	0	2
always intense (%)	8	5	5	5	2	2	2	0	0	0	0	5	2	0	2	0
M	0.97	0.72	0.69	0.64	0.56	0.47	0.44	0.39	0.38	0.38	0.34	0.31	0.27	0.19	0.16	0.13
SD	1.10	1.02	0.96	0.98	0.79	0.73	0.71	0.55	0.55	0.60	0.51	0.91	0.70	0.43	0.57	0.38

The questionnaire assesses how frequently (never 0, occasionally 1, and always 2) and how intensely (moderate 1, intense 2) a symptom occurs. Symptoms are ordered with regard to severity from left to right with "headache" being the most frequent symptom and "colored halos around objects" the least frequent one. We show the ratings for each symptom in percent. Below the mean single symptom score of each symptom is given; averaged over all participants. The single symptom score of a symptom is calculated as the product of frequency and intensity.

followed by *excessive blinking* ($M = 0.72, SD = 1.02$), *dryness* ($M = 0.69, SD = 0.96$), and *blurred vision* ($M = 0.64, SD = 0.98$) (see Table 1). We found a mean single symptom score of $M = 0.44$ ($SD = 0.23$).

We asked respondents whether they had ever experienced a symptom that was not listed. We received three answers for this question: "strong eye strain", "sweat and VR headset weight pressing on my face", and "tiredness after several hours (heavy eyelids)".

3.3.5 Coping Strategies. In this section, we asked participants about their knowledge and usage of coping strategies to alleviate DES symptoms. We listed six commonly known coping strategies and asked respondents whether they knew and used them (see Figure 3 left). In general, 66% of the respondents stated to use coping strategies during VR usage. Total 11% indicated to apply them two or three times an hour, 31% once an hour, and 58% less than once an hour (see Figure 3 right). There were some cases where participants knew about a coping strategy but did not use it. Reasons for this are that participants indicated to forget about them (46%) or considered applying them as too interruptive (23%) or effortful (8%). Total 34% stated that it would not deem necessary. Participants had the opportunity to indicate additional coping strategies that they used. Here, participants mentioned the following strategies: "I washed them [the eyes] with water"/"Splash my eyes with cold water", "Drink some water", "Covering one eye for a while to fix depth perception.", "Just taking the headset off for 5 minutes does the job for me", and "for about 10 seconds I Focus on a nearest thing (mostly my fingers raised to eye level) and after 10 seconds, I look far away and Focus on a distant thing (a top of a tree most often)".

None of the participants mentioned using an eye exercise or eye training mobile app. When asked about further comments regarding strategies to prevent or mitigate visual discomfort, we received nine answers. One person asked for apps that could be used (P70), and one person mentioned that they did not have problems and therefore no need to apply coping strategies (P60). Two respondents indicated that they did not know that strategies existed and that they usually just closed their eyes (P37) or stopped the VR session to do nothing for a while (P47). P23 mentioned doing them more often "now that I learned more of them". Two respondents mentioned that they thought that "[eye exercises] ought to be recommended in the VR software" (P62), and that

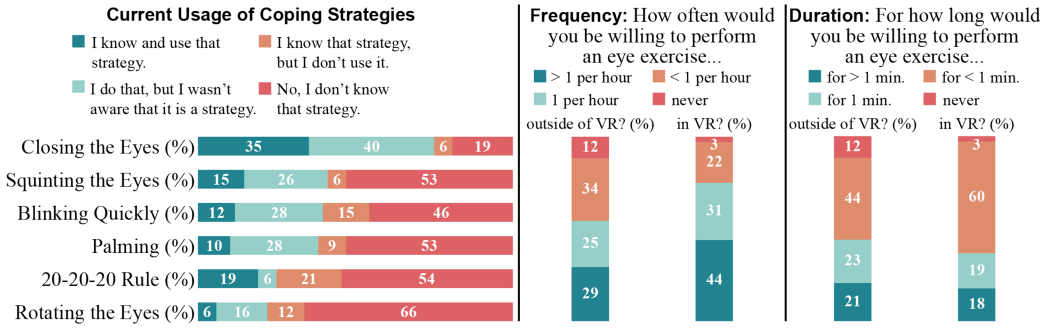


Fig. 3. Left: respondents' answers to whether they knew and already used any of the presented coping strategies. Right: frequency and duration preferences on how respondents would prefer to use coping strategies without VR-HMD and integrated into a VR-HMD.

"[eye exercises] must be more publicized in order to increase awareness" (P31). Two respondents mentioned that they used "splash of water or lubricant eye drops" (P69) or that they "instinctively did things to cure our eyes" (P51).

3.3.6 Usage of Coping Strategies. Lastly, we asked participants about their willingness to perform coping strategies to reduce DES symptoms during VR usage. When asked if users would take off the headset to perform eye exercises, 75% agreed. When asked about performing them in VR, the agreement was at 80%, and 89% agreed to perform eye exercises that were integrated implicitly into VR (e.g., in loading screens). Preferred frequency and duration of eye exercises are listed in Figure 3 on the right.

When asked for final comments, we received the following answers. One participant mentioned that not everybody experiences DES the same and therefore "the feature should be optional and the length/frequency should be customizable" (P37). P23 indicated their willingness to do strategies during loading screens in VR: "it would create a better experience even for the player". P62 stated that "more games and VR sets should use them [coping strategies], it is important to not scare people away from the technology". P60 even wrote that coping strategies "should be [of] knowledge for everyone who uses VR".

3.4 Discussion

The results of our online survey confirm that DES is widespread and a severe problem to VR users, with half of our survey participants taking off the HMD at least once an hour due to symptoms of DES. Some (14%) even continue to experience symptoms for an hour or more after usage. The results of the CVS-Q further support these findings, indicating that half of the respondents can be considered to suffer from CVS (with 53% of the respondents having a score ≥ 6). We also found notable differences across users: While some mentioned that discomfort occurred very shortly (within 15 minutes) of using the device, others experienced it only after three or more hours of usage. For example, P27 indicated that they got a headache after using the device "longer than [they] usually do". These results suggest that prolonged use should not be defined as an absolute time value, but individually and adapted to a user's habits, since—as stated by P37—"not everyone experiences the discomfort the same".

If DES occurred, respondents stated to interrupt their experience most often either by discontinuing device use or by closing their eyes for a while. About half of the participants knew coping

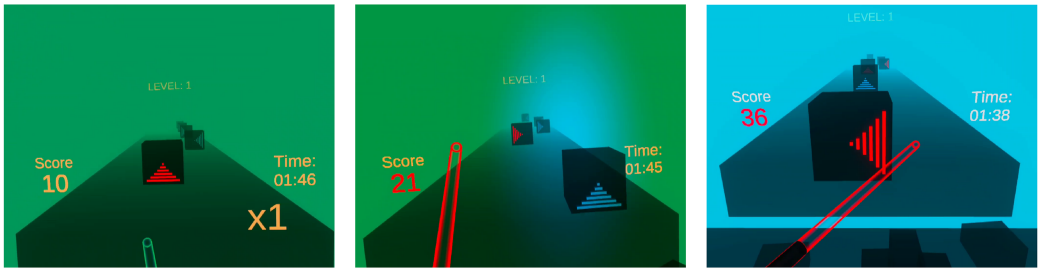
strategies to mitigate eye strain. Most respondents who experienced DES symptoms applied their strategies spontaneously (e.g., "I didn't know that strategies existed, I tend to just close or rub my eyes [or] take breaks from the screen", P47). As a potential problem, we observed that respondents forget about strategies and ignore DES when immersed in the experience. These findings underline the strong need to develop solution techniques that can be integrated into VR-HMDs. This could be implemented as a constant overlay present during the experience (e.g., blue light filtering) or could be integrated as discrete, automated breaks during the experience (e.g., eye exercises). We found strong support for the idea of these device-integrated solutions among participants, with 80% being willing to perform eye exercises integrated into VR-HMDs, and even more (89%) to use them embedded in the experience (e.g., in loading screens or similar). Several participants also made corresponding comments, e.g., "I think they ought to be recommended in the VR software" (P62).

Regarding the specific symptoms that participants experienced, six symptoms had a single symptom score greater than the mean single symptom. These were *headache*, *excessive blinking*, *dryness*, *blurred vision*, *increased sensitivity to light*, and *eye redness* (see Table 1). Three of these symptoms (*excessive blinking*, *dryness*, and *eye redness*) are related to external symptoms, according to Sheedy et al. [85]. Internal symptoms can be caused by accommodative and convergence stress [86]. Therefore, it is assumed that internal symptoms might become especially significant in stereoscopic displays [103]. Our results support this assumption, as *headache*, which is an internal symptom, is the most severe symptom. Furthermore, *eye pain*, which is also an internal symptom, has a score equal to the mean symptom score, and is therefore also relevant to users. Lastly, the symptom *double vision* was not very important to the users, although more than 40% of the survey respondents did not adjust the IPD of their VR-HMD. Therefore, we assume that other effects were stronger in causing DES symptoms than the IPD adjustment.

The first conclusion we derive from the survey is that DES occurs frequently and affects more than 50% of users regularly (at least once an hour of using a VR-HMD). Secondly, DES causes most affected users to interrupt their VR experience or discontinue using the device completely. Users also apply individual coping strategies, the effectiveness of which has not been determined. These results underline the need to develop novel solutions to DES in VR that avoid interruptions and are integrated into VR-HMD experiences. This was accurately summarized by P62, who stated, "More games and VR sets should use them [eye exercises]. It is important not to scare people away from the technology like we did with non-LED screens".

4 ADDRESSING DES IN VR-HMDS

The results of our online survey showed that DES symptoms are widespread among VR-HMD users and that alleviation approaches are missing. In particular, respondents' statements highlighted the need for integrated treatment methods that they can use without interrupting device use. In this section, we, therefore, present our investigation of two types of device-integrated treatment methods. First, we investigated the blue light filtering approach, which is widely used with conventional digital devices to alleviate DES symptoms, and has been integrated into VR-HMDs by some manufacturers (e.g., as "night mode" in the Oculus Quest[2]). Secondly, we investigated eye exercises, i.e., short visual tasks that trigger specific eye movements, aiming to relax the eye muscles and reduce temporary symptoms of DES. Both treatment approaches can be integrated into the device use, either by applying them constantly (blue light filtering) or by providing discrete breaks in the experience to perform them (eye exercises). We conducted three user studies to evaluate both approaches in detail.



(a) A screenshot of the *LightSaber* game with the **global filter**.

(b) A screenshot of the *LightSaber* game with the **peripheral filter**.

(c) A screenshot of the *LightSaber* game with **no filter**.

Fig. 4. Three screenshots of the custom VR game *LightSaber* that we developed and applied as the study task in the blue light study.

4.1 Blue Light Filtering to Address DES in VR-HMDs (User Study 1)

Blue light filtering is broadly distributed as a method to protect users' eyes during digital device usage (e.g., on mobile phones[4]). Already, some VR-HMD manufacturers offer blue light filters integrated as "night mode", e.g., the Oculus Quest[2] or HTC Vive.[5] However, studies with conventional computer displays could not find a positive effect of blue light filters on DES symptoms [69, 76], and an empirical investigation of the effect of blue light filters in VR-HMDs is missing. We designed and conducted a user study to investigate the potential effect of blue light filters on DES symptoms in VR-HMDs.

If effective, blue light filters would constitute an easy-to-integrate solution for VR-HMD devices and experiences. However, the changes in colors and reduced brightness could cause adverse effects on usability metrics, such as enjoyment, visual appeal, presence, and obtrusiveness. To integrate these aspects into our study design, we implemented two versions of blue light filtering and compared these to a control condition. In the first version, a filter was applied *globally* to the virtual image, similar to blue light filters in conventional smartphone or laptop displays. The filter was implemented to block 60% of the emitted blue light (see Figure 4(a)). We determined this percentage in a set of pre-study experiments described in appendix Section B.1. Furthermore, we investigated a second, *peripheral* version of blue light filtering, where a filter is only applied to the users' peripheral field of view. This second version constitutes a trade-off between the potential positive effects on DES symptoms and negative effects on usability metrics. The visual system is only sensitive to color in the foveal part of the field of view [96]. Therefore, we hypothesized that when using a peripheral filter, users would still benefit from perceiving the colors naturally in the foveal part while getting the positive effects of a lower amount of blue light that enters the eye from the periphery. We implemented the filter as a radial vignette (25°) centered at the user's head pointer, from which the intensity of blocked blue light increased exponentially from 20% to 60% at the periphery (see Figure 4(b)). To determine the specific properties of the peripheral filter, we conducted three pre-study experiments that are described in detail in the appendix Section B.1.

We expected that—if blue light filters affect DES symptoms—a globally applied filter would result in lower DES symptoms than a peripheral filter. However, the peripheral filter should still reduce symptoms in comparison to a control condition. On the other hand, for the usability metrics (enjoyment, visual appeal, presence, and obtrusiveness), we expected that scores would be higher for

[4]Samsung: https://www.samsung.com/au/support/mobile-devices/about-the-blue-light-filter/, last retrieved: April 29, 2021.
[5]HTC Corporation: https://www.htc.com/us/support/htc-10/howto/night-mode.html, last retrieved: April 30, 2021.

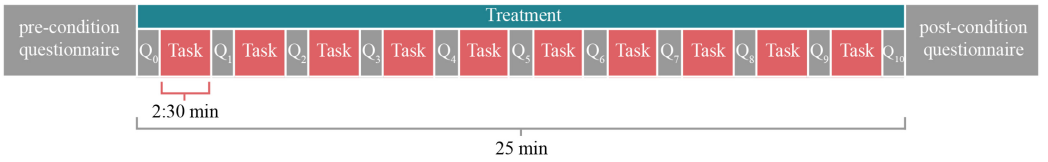


Fig. 5. The procedure for the blue light study. Participants answered each a complete questionnaire before and after each condition. During each condition, participants indicated their current perception of symptoms by answering intermediate single-item questions on DES, simulator sickness, and ergonomic symptoms, indicated by Q_0–Q_{10}. This procedure was the same for all three conditions, which were conducted on three different days. During one condition, one type of treatment was active, i.e., a global filter, a peripheral filter, or no filter.

the peripheral filter than for the global filter, as a peripheral filter would cause fewer color changes in the environment. In summary, the following hypotheses guided the study:

H1: With a **global** blue light filter, participants report less severe DES symptoms than with **no filter**.
H2: With a **peripheral** blue light filter, participants report less severe DES symptoms than with **no filter**.
H3: With a **global** blue light filter, participants report less severe DES symptoms than with a **peripheral** filter.
H4: With a **global** blue light filter, participants report lower usability scores than with **no filter**.
H5: With a **peripheral** blue light filter, participants report lower usability scores than with **no filter**.
H6: With a **global** blue light filter, participants report lower usability scores than with a **peripheral** filter.

4.1.1 Study Design. We implemented the study as a repeated-measures design with one independent variable **treatment** with the three levels *global filter*, *peripheral filter*, and *no filter* (control condition). The order of conditions (three) was counterbalanced (six groups in total), and we distributed participants randomly to the six groups. We recruited 28 participants and therefore had imbalanced groups. The study software distributed the participants to four groups consisting of four participants (two female, two male) and two groups consisting of six participants (three female, three male).

4.1.2 Measures. We assessed DES symptoms subjectively using self-report measures. Symptoms were assessed before and after each condition, using a complete questionnaire (see pre/post-condition questionnaire in Figure 5). The complete DES questionnaire is based on Hirzle et al.'s analysis of symptoms in VR-HMDs [39]. It covers 21 symptoms on a 7-point scale, reaching from *nothing at all* to *very severe*. In particular, we instructed participants as follows for each of the items in the questionnaire: "click how strong your perception of each symptom is right now." In addition to the complete pre and post-condition questionnaire, we assessed a single-item of DES 11-times during the study, integrated as intermediate question (indicated as Q_0–Q_{10} in Figure 5). It was presented as follows: "Please indicate how strong your perception of DES symptoms is right now." (7-point scale, ranging from *nothing at all* to *very severe*).

Furthermore, we assessed symptoms of simulator sickness and symptoms caused by the headset's ergonomics with complete questionnaires after each condition, also based on Hirzle et al.'s analysis [39]. While we asked about simulator sickness symptoms before and after each condition,

we assessed ergonomic symptoms only after each condition. Symptoms of simulator sickness cover the 16 items of the simulator sickness questionnaire [50]. We assessed ergonomic symptoms with 18 items, covering the six comfort rating scales of wearable devices [55], six statements about the attachment of wearable devices [16]. Furthermore, we assessed the intensity of discomfort on different regions of the head and face with nine items on a Borg CR 10 scale [14]. Like the questions on DES, we asked participants to indicate how strong their perception of each symptom was right now on a 7-point scale, reaching from *nothing at all* to *very severe*. During the study, we integrated a single-item version of each symptom group (simulator sickness and ergonomic symptoms), similar to the single-item intermediate question on DES. The complete questionnaires are presented in the appendix Section B.3.

In addition, after each condition, we assessed enjoyment, visual appeal, presence, and obtrusiveness, and polled for whether participants adjusted the interpupillary distance of the VR-HMD and whether they wore vision aids during the study. To assess *enjoyment* we used the corresponding sub scale of the **Player Experience Inventory** (**PXI**) by Vanden Abeele et al. [1] (five items on a 7-point Likert scale, ranging from *strongly disagree* to *strongly agree*). To measure *visual appeal* we used three items of the PXI's audiovisual appeal sub scale that are also assessed on a 7-point Likert scale, ranging from *strongly disagree* to *strongly agree*. For the assessment of *presence*, we used the **igroup presence questionnaire** (**IPQ**) [83] (14 items on a 7-point scale). For the measurement of *obtrusiveness*, we assessed five items from our pre-study experiment on a 7-point Likert scale, ranging from *strongly disagree* to *strongly agree*. These items covered whether participants felt that the visuals of the games influenced their experience and whether they noticed any changes in the visuals. See appendix Section B.4 for the complete questionnaires.

4.1.3 Participants. To determine the sample size, we used the G*Power software.[6] We conducted the following type of power analysis: "F tests", "ANOVA: Repeated measures, within factors", "A priori: Compute required sample size—given α, power, and effect size". The parameters we entered were: "effect size: 0.25", "αerr prob: 0.05", "Power (1-βerr prob): 0.8", "number of groups: 1", "number of measurements: 3", "corr among rep measures: 0.5", and "nonsphericity correction η: 1". We assumed a medium effect size of $f = 0.25$ based on previous research on blue light filters [76]. Furthermore, as blue light filtering is applied by many consumers of digital devices, we expected an at least medium sized effect. This analysis resulted in a total sample size of 28 participants.

Consequently, we recruited 28 participants (14 female, 14 male). To ensure that participants paid attention to the task and questionnaires, we integrated three attention checks into the pre and post-condition questionnaires and three attention checks into the intermediate questions for each condition. The attention checks adhere to the survey platform's guidelines on fair attention checks.[7] No participant missed more than one attention check in an intermediate question or the pre and post-condition questionnaires. Therefore, we included the data of all 28 participants in the final analysis.

4.1.4 Study Task and Apparatus. As the study task, we implemented a customVR game *LightSaber* based on the popular VR game BeatSaber.[8] In the game, the players have a red and a blue saber that they use to cut cubes, which are spawned according to the music's beat and that fly toward them. The cubes are highlighted in either red or blue and can only be cut with

[6]G*Power Software: https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html, v 3.1.9.7, last retrieved: March 24, 2021.
[7]Prolific attention checks: https://researcher-help.prolific.co/hc/en-gb/articles/360009223553-Using-attention-checks-as-a-measure-of-data-quality, last retrieved: April 8, 2021.
[8]BEAT GAMES: https://beatsaber.com/, last retrieved: March 23, 2021.

the respective saber. Also, players have to cut the cubes in the correct direction indicated on each cube with an arrow. The goal of the game is to cut all the cubes with the correct saber at the right time (see Figure 4 for three screenshots of the game). The cubes were spawned from eight random positions (top, bottom, left, right, upper left, upper right, lower left, and lower right) and four random rotations (0°, 90°, 180°, and 270°, see appendix Section B.5). The *LightSaber* game consisted of 10 levels, increasing in difficulty. One level took 2:30 minutes. In the first two levels, cubes only spawned from the left and the right position. In the third and the fourth levels, cubes additionally spawned from the top and from the bottom position. For levels five to eight, cubes spawned from the upper left, upper right, lower left, and lower right positions, and in the final two levels (nine and 10), cubes spawned from all eight positions. We used 10 different songs that increased in beats per minute (145, 148, 148, 178, 184, 190, 210, 214, and 220). A list of the songs is given in the appendix Section B.6.

We pursued two goals with the design of the study environment. First, it had to be designed in a way to induce symptoms of DES in a relatively short amount of time. Secondly, it had to be designed in a way that blue light filtering would have an effect. If the game was designed with colors with a low amount of blue light, blue light filtering would not have a large effect. Therefore, we designed the game environment with a large amount of blue light. Furthermore, we styled game elements to induce DES symptoms by using very bright and glowing colors to design the sabers and the cubes. We implemented the study using Unity 3D[9] and the Oculus Quest[2]. A video of the game, including the two blue light filters, can be found in the supplemental material.

4.1.5 Procedure. We conducted the study as an online study, recruiting participants via Prolific.[3] In a registration survey, we polled whether participants owned an Oculus Quest (version one or version two) and whether they would like to participate in a three-day study. We only sent invitations to the main study to participants who registered for participation. We split the main study into three parts that participants had to conduct on three different days. We instructed participants to conduct the study at the same time every day to exclude external discomfort factors caused by different times of the day. Participants only received an invitation to the second and third day of the study after successfully completing the first and second day, respectively. We only sent an invitation to the next day of the study 12 hours after participants had completed the previous day to ensure that participants did not play the application twice on the same day. We scheduled the study to take three hours in total, and participants received a reward of £27 in total. We did not record any personal information from the participants, but the survey platform provided the following additional data: age, sex, country of birth, country of current residence, employment status, first language, nationality, and student status.

We clearly stated and explained that participants had to download a study software provided by the researchers. After the participants provided informed consent, we introduced them to the study procedure. They then filled out the two pre-condition questionnaires on DES and simulator sickness. Only on the first day, we presented a guide on downloading and installing the study software to their device. Afterward, we explained the three questions that participants had to answer during the application, once in the beginning and then after completing each level. The participants answered the questions inside the same VR environment, using the same game controls as during the game, as this was shown to preserve presence in game environments [35]. Participants could refer to a board to their right, which explained the three symptom classes and the single symptoms for each category. After giving a study ID to participants, the application started with introducing the *LightSaber* controls and a tutorial on cutting the cubes in the correct direction and

[9]Unity Technologies: https://unity.com/, last retrieved: March 24, 2021.

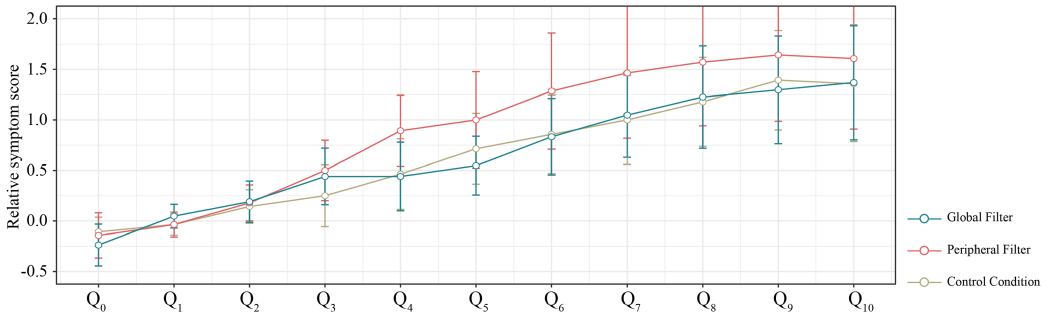


Fig. 6. The results of the intermediate single-item DES statement "Please indicate how strong your perception of digital eye strain symptoms is right now" (7-point scale, ranging from *nothing at all* to *very severe*) that was polled before the first level (Q_0) and after each level (Q_1–Q_{10}) in each condition. We show the relative symptom score averaged over all participants for each condition. Bars represent standard deviation.

Table 2. The Four Symptom Means Calculated for the Blue Light Study Based on Sheedy et al.'s Symptoms Classification [86]

	M_{ex}	M_{in}	M_{vr}
M_{all}	burning irritation dryness	strain headache	sensitivity to bright light difficulty focusing blurred vision double vision

speed. After playing all 10 levels of the application, the study data were automatically uploaded to a secure server of the university, and participants were referred back to the online survey. They then answered the three post-experiment questionnaires and the questions about the VR experience.

4.1.6 Results. Our analysis will focus on the intermediate question and pre and post-condition questionnaires about DES, which is the primary concern of this work. The results of the measures of simulator sickness and ergonomic symptoms are not within the scope of this work.

Intermediate Question Results. In this section, we report the findings about the single-item DES measure that was polled 11 times during each condition, indicated as Q_0–Q_{10} in Figure 5. We calculated a relative symptom score by subtracting the mean of the first three values as baseline ($mean(Q_0, Q_1, Q_2)$) from all values Q_0–Q_{10}. The rationale and exact calculations of this baseline are given in Section B.2. The time course of the within-condition DES values averaged over all participants for each condition is shown in Figure 6.

To determine potential effects of **treatment** and **time point** on DES, we conducted a two-factorial analysis of variances. As normality tests revealed that the answers significantly differed from a normal distribution for more than one group, we used a non-parametric test for the analysis of variances (see appendix Section B.7.1 Table 16 for the results of the Shapiro–Wilk normality tests). The nparLD R-package provides an ANOVA-type non-parametric test for the analysis of variances [67].

We calculated a two-factor non-parametric ANOVA with the two factors **treatment** with three levels (*global filter, peripheral filter*, and *control condition*) and **time** with 11 levels ($Q_0 - Q_{10}$). In the following, we report ANOVA-type test statistics and *r-equivalent* effect sizes based on Rosenthal

Table 3. Friedman Test Results and Mean Values of the Four Symptom Means Overall Symptom Mean M_{all}, Symptom Mean of the External Symptoms M_{ex}, Symptom Mean of the Internal Symptoms M_{in}, and Symptoms Mean of the VR-specific Symptoms M_{vr}, Shown for Each Condition of the Blue Light Study

	M_{all}	M_{ex}	M_{in}	M_{vr}
Friedman test results	$\chi^2(2) = 0.44, p = .80$ $W = .0079$	$\chi^2(2) = 1.23, p = .54$ $W = .0219$	$\chi^2(2) = 0.08, p = .96$ $W = .0013$	$\chi^2(2) = 0.19, p = .91$ $W = .0034$
Global filter	$M = 0.32, SD = 0.44$	$M = 0.29, SD = 0.70$	$M = 0.71, SD = 1.02$	$M = 0.20, SD = 0.40$
Peripheral filter	$M = 0.42, SD = 0.61$	$M = 0.54, SD = 1.04$	$M = 0.68, SD = 0.89$	$M = 0.21, SD = 0.44$
Control condition	$M = 0.27, SD = 0.39$	$M = 0.35, SD = 0.70$	$M = 0.63, SD = 0.89$	$M = 0.07, SD = 0.35$

Friedman tests did not reveal an effect of treatment on the symptom means. We report Kendall's W effect sizes.

and Rubin's method [82].[10] The test indicated a moderate effect of **time** ($F(2) = 51.97, p < .01, r = .53$) on relative DES values. Pairwise comparisons using post-hoc Dunn's tests with Bonferroni correction revealed statistically significant differences between the majority of time points. However, we did not find a statistically significant effect of **treatment** ($F(2) = 2.29, p = .11, r = .21$) nor an interaction effect of **time** and **treatment** ($F(1) = 7.58, p = .18, r = .23$).

In summary, the level of DES significantly increased over the exposure time of 25 minutes in all three conditions, and we could not find a statistically significant effect of treatment on the DES levels between the conditions.

Pre/Post-Condition DES Results. For the DES questionnaires, we covered four aspects of eye strain represented by four symptom means: M_{all}, M_{ex}, M_{in}, M_{vr} (see Table 2). We analyzed the potential effects of **treatment** (*global filter*, *peripheral filter*, and *control condition*) on the relative changes[11] of DES symptoms before and after each condition. We used a Friedman test, as the variables significantly differed from a normal distribution (see appendix Section B.7.1 Table 17 for the results of the Shapiro–Wilk normality tests). The results did not indicate a statistically significant difference between the four symptom means of the treatment methods (see Table 3 for statistical results).

In summary, we could not observe an effect of the treatment methods on the DES levels, not in the intermediate questions nor in the pre/post questionnaires.

We further analyzed the internal consistency of each of the symptom means with Cronbach's alpha. We found acceptable values for M_{all} ($\alpha = 0.77$) and M_{ex} ($\alpha = 0.77$), but a slightly low values for M_{in} ($\alpha = 0.67$), and a poor value for M_{vr} ($\alpha = 0.51$). Therefore, the results for M_{vr} must be considered with caution.

Pre/Post-Condition Usability Results. We obtained values for the *enjoyment*, *presence*, *visual appeal*, and *obtrusiveness* scales by averaging their 7-point scores (results for the Shapiro–wilk normality tests are shown in the appendix Section B.7.1 Table 17). Internal consistency, measured with Cronbach's alpha, was good for all four scales: *visual appeal* ($\alpha = 0.92$), *enjoyment* ($\alpha = 0.97$), *presence* ($\alpha = 0.87$), and *obtrusiveness* ($\alpha = 0.71$). Statistical results and mean values of these variables are shown in Table 4.

Using a Friedman test, we did not find statistically significant differences between the means for visual appeal, enjoyment, presence, or obtrusiveness.

[10]This effect size is calculated based on the $p\text{-}value$, the sample size, and the number of conditions. It is designed for situations where, for instance, "non-parametric procedures were used for which there are not currently accepted effect size indicators" [82].

[11]We calculated relative values by subtracting the value of the pre-condition questionnaire from the value of the post-condition questionnaire for each symptom separately.

Table 4. Friedman Test Results and Mean Values of the Usability Metrics Visual Appeal M_{VA}, Enjoyment M_E, Presence M_P, and Obtrusiveness M_O

	M_{VA}	M_E	M_P	M_O
Friedman test results	$\chi^2(2) = 1.51, p = .47$ $W = .0269$	$\chi^2(2) = 1.21, p = .55$ $W = .0216$	$\chi^2(2) = 3.16, p = .21$ $W = .0564$	$\chi^2(2) = 2.8, p = .25$ $W = .0500$
Global filter	$M = 3.37, SD = 1.42$	$M = 3.99, SD = 1.31$	$M = -0.18, SD = 0.87$	$M = 3.16, SD = 0.94$
Peripheral filter	$M = 3.58, SD = 1.40$	$M = 4.18, SD = 1.27$	$M = -0.06, SD = 0.87$	$M = 3.01, SD = 0.94$
Control condition	$M = 3.63, SD = 1.39$	$M = 4.17, SD = 1.29$	$M = -0.03, SD = 0.76$	$M = 2.78, SD = 0.73$

For each mean value, we conducted a Friedman test to reveal a potential effect of treatment. We report Kendall's W effect sizes.

4.1.7 Discussion. Despite the widespread popularity of blue light filters for digital devices, we could not find statistically significant effects of blue light filters in VR-HMDs, not in the intermediate single-item measured during the application, nor in the more detailed complete DES questionnaire. Therefore, we reject hypothesis **H1**. We did also not find an effect for the peripheral filter compared to the no filter condition during the application. In contrast to the expected positive effect of the peripheral filter, descriptive results suggested that it caused even more severe DES values than when no filter was applied. Therefore, we also reject hypotheses **H2** and **H3**. This effect could be explained by the implementation of the filter based on the head-point of the participants. As we used a headset that did not have eye tracking integrated, we could not adapt the filter to the gaze point of the participants, but only to their head rotation. Therefore, when participants moved their eyes but did not move their head, the filter would not change, but participants could be able to detect the filter. Although the subjective responses indicate that participants did not detect the filter, we assume that the slight movement of the eyes might have detected the color changes in the periphery. These slight changes might have influenced the perception of the filter. Similar to an afterimage, looking at a blue light filter for a specific time period causes fatigue of the peripheral cells. When looking at an image without filter directly afterward, it, therefore, appears to be even brighter than before. This effect might have been initiated by the peripheral filter and might therefore have caused an increase in DES values.

We expected that when using a blue light filter enjoyment, presence, visual appeal, and obtrusiveness would be affected negatively. However, we could not observe any differences for these values. Therefore, we reject hypotheses **H4–H6**. These results indicate that participants were not only not disturbed by the filters, but that they did not even realize that a filter was active. One explanation could be that participants were too immersed in the VR experience and too concentrated on the game that they would notice a filter, although we applied strong filter values of 60%. Another explanation could be that the effect that the filter caused (both in DES symptoms and in usability metrics) was very small and therefore overshadowed by effects of the game. Lastly, the experiment is naturally limited by the VR application. Although we tried to create a DES-inducing environment, and investigated the settings of the blue light filters with several pre-study experiments, the absence of an effect could be attributed to our custom implementation.

We conclude that if there is an effect of blue light filters on DES symptoms in VR-HMDs, it can only be a small effect, as we were not able to detect a statistically significant effect in our study. This is despite the fact that we used an extra-strain-causing environment, a comparable exposure time to previous DES studies [69, 81], well-established measures in the investigations of DES symptoms [39, 86], and an adequate sample size to detect a medium effect, as indicated by the power analysis. With all the precautions that we discussed, we conclude that applying a blue light filter does not help significantly to reduce DES symptoms. Our results are in agreement with three recent studies,

where no effect of the blue light filter condition could be found on DES symptoms in comparison to a control condition in conventional displays [69, 76, 81].

4.2 Applying Eye Exercises to Address DES in VR-HMDs

The second approach that we investigated to address DES in VR-HMDs is eye exercises. The survey results reveal that respondents are positive about using coping strategies, including eye exercises, in general, and could envision using them regularly to reduce DES symptoms in VR- HMDs. To empirically investigate the effect of eye exercises, we designed and evaluated a set of them in two user studies. We designed the eye exercises as short visual tasks that trigger specific eye movements, aiming to relax the eye muscles and reduce temporary symptoms of DES. In the design of the eye exercises, we focused on their potential to be integrated into a VR application, i.e., to address DES without requiring users to take the HMD off. As some survey respondents (e.g., P37) mentioned closing their eyes when DES symptoms occur, we integrated *closing the eyes* as a special type of eye exercise.

In the following, we first detail the design of the eye exercises. We then present the results of the two user studies in which we investigated their effectiveness.

4.2.1 On the Design of Eye Exercises. For the design of eye exercises, we focused on the prevalence of symptoms as indicated by the survey respondents. The most prevalent symptoms, here listed with classification as internal (I) or external (E) symptom, based on Sheedy et al.'s [85] analysis (an (X) indicates that the specific symptom is not classified as an external or internal symptom), are headache (I), excessive blinking (E), dryness (E), blurred vision (X), increased sensitivity to light (X), and eye redness (E). For the design of the eye exercises, we started by identifying causes that have been linked to these symptoms [24]. We then designed eye exercises inspired by commercial vision therapy applications[12] that activate specific eye muscles and, in turn, aim at reducing symptoms of DES. Each eye exercise was designed to take 30 seconds based on the online survey results, where 60% of the respondents stated to be willing to perform coping strategies for up to 1 minute (see Figure 3 right). In the following, we detail the design of eye exercises for external and internal symptoms.

Eye Exercises to Address External Symptoms. External symptoms are linked to dry eye syndrome, which is strongly connected to a decrease in blink rate [12, 79]. The literature further suggests that the completeness of blinks is decisive, with incomplete blinks causing the eye to dry out faster [23, 73]. To counter dry eye, we designed **E1** —a blinking task for which users have to perform voluntary blinks for 30 seconds with closing the eyelids fully to prevent incomplete blinks (see E1 in Figure 7). Users are instructed by a pictogram of open/closed eyes and an audio cue to perform the blinks (see E1 in Figure 8). The eye exercise is implemented as one voluntary blink per second for seconds 0–10 and one blink per 0.75 seconds for seconds 11–20, i.e., the blink rate was slightly increased after the first 10 seconds.

Eye Exercises to Address Internal Symptoms. While the link of external symptoms to dry eye is clearly established, the causes of internal symptoms are more versatile. Glare and lighting are known to cause DES with desktop monitors [79]. Increased sensitivity to light and headaches are most probably caused by an increased amount of glare in VR-HMDs, where the eye-screen distance is reduced, and the field of view is largely occupied [79]. This might result in the *tension of pupillary muscles* that have to constantly adapt to very bright lighting conditions in VR. **E2** aims at relaxing the pupillary and ciliary muscles that control the pupil and the eye lens by contraction and

[12] https://play.google.com/store/apps/details?id=com.eyeexamtest.eyecareplus&hl=en_US, last retrieved: March 31, 2021

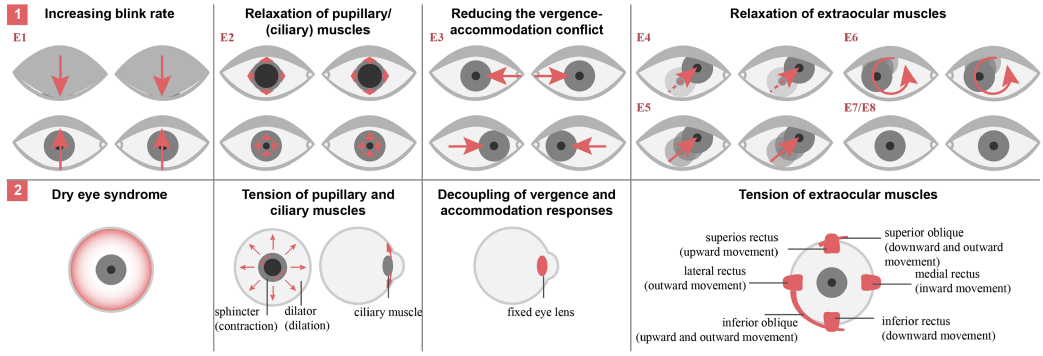


Fig. 7. We show the relation between each eye exercise (1) and symptom that we aim at addressing with the respective exercise (2). **E1** is designed to address the dry eye syndrome by increasing blink rate and the completeness of blinks. **E2** is designed to reduce tensions in the pupillary and ciliary muscles by changing brightness level of the background and not providing a visual cue. **E3** requires the eyes to shift focus to a far distance, decreasing the VAC. **E4**–**E8** are designed to release tension of the extraocular muscles, either by inducing motion **E4**–**E6** or by reducing motion **E7**–**E8**.

Fig. 8. Screenshots of the custom VR eye exercise application: (a) the *WordSearch Puzzle* implemented as a hemisphere of letters, (b) point of view of participants during the study task, this was same for the control condition, (c) + E1–E8: instructions for the eye movements (arrows depicted here to indicate animations within the instructions), and closing the eyes, **E1:** periodic blinking, **E2:** pupillary light reflex (with changing background), **E3:** fixation shift, **E4:** saccades, **E5:** smooth pursuit, **E6:** rolling eyeball into one direction, **E7:** static fixation, and **E8:** loose focus with open eyes. A video of the eye exercises is given in the supplemental material.

relaxation [7]. We trigger the user's pupils to react to slow lighting changes and, hence, contracting and de-contracting the underlying muscles. In **E2**, we change the brightness of the background from bright to dark, causing the pupil to slowly adapt to these changes (see Figure 7 E2).

The *vergence-accommodation conflict* in HMDs occurs as a result of non-corresponding depth cues for vergence and accommodation responses. While the binocular disparity in VR triggers vergence responses, accommodation responses are decoupled, and the eye lens focus is limited to only one focal plane. As such, the ciliary muscle that controls the eye lens keeps being contracted to bend the lens to the proper focal distance. This can result in symptoms like headache, blurred vision, or difficulty focusing [42, 52]. As accommodation is set to infinity in common VR-HMDs, a potential solution to minimize the VAC is to display objects at a far distance. With **E3**, we designed an eye exercise that triggers a dynamic fixation, reaching from close to far and as such trigger vergence responses to counteract the high demand of near vision in HMDs. At its farthest point, the VAC should be minimal and, therefore, this exercise could temporarily reduce symptoms thereof. Users were asked to keep fixating a virtual sphere that changes its position from 2.6 Unity units distance to 0.4 Unity units distance two times in 30 seconds.

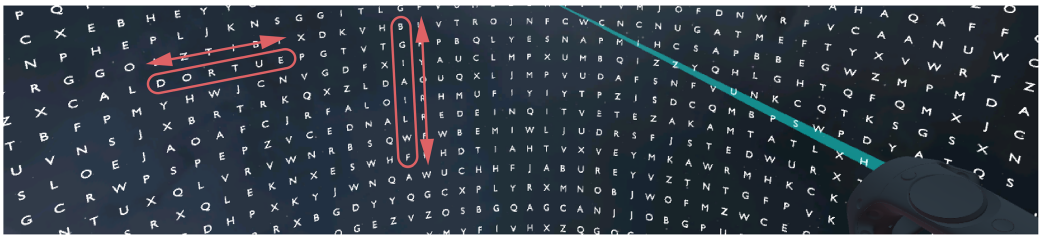


Fig. 9. We implemented a custom VR experience *WordSearch Puzzle* that served as the task in both user studies on eye exercises. Participants' task was to find hidden words that were arranged randomly in four directions.

One cause for headache, difficulty focusing, and general discomfort is tensions of the extraocular eye muscles produced by consistent focusing of the eyes on interactive elements in the foveal field of view and the high demand of near vision in VR-HMDs. The extraocular muscles are those that move the eyeball [5, p. 69]. To release the eyes from a possible tension, we designed three types of dynamic exercises that activate these muscles by moving the eye into different directions and applying different rotations. In **E4,** a sphere is presented that jumps to random locations in the user's inner and outer field of view. Following the sphere with the eyes causes saccades and fixations, resulting in contraction and relaxing of the extraocular muscles (see Figure 7 E4). To perform this exercise, users have to fixate a virtual sphere that randomly jumps at different positions in the virtual space within the field of view. We also provide a version of this exercise that relies on the eye movement smooth pursuit (**E5**). Here, users have to fixate a virtual sphere that floats at a random pattern in the virtual space within the field of view. **E6** involves moving the eyes into one direction for 15 seconds each (see Figure 7 **E6**). The user is instructed by an arrow that indicates the direction of rotation.

Lastly, as the task we chose to induce symptoms was a reading task (we describe the task in detail in Section 4.2.2), we wanted to offer two counter eye exercises that let the eyes rest for some time, one with and one without a stimulus. These are **E7** and **E8**. In the first one, a black background is shown, and participants are simply asked to look ahead. In the second one, a stimulus (virtual sphere at 1.5 Unity units) is shown that participants should concentrate their gaze upon for the pre-defined time.

Lastly, we consider *closing the eyes* as a special case of eye exercise, as it does not provide a visual stimulus, but simply asks the users to close their eyes for 30 seconds (see (c)) in Figure 8).

Implementation of Eye Exercises. For implementing the eye movements as VR experiences, we developed a framework that allows developers to combine an optional number out of the eight eye exercises plus *closing the eyes* into specific sessions. Each exercise starts with a one-sentence instruction, which the user has to confirm by clicking on a start button on the controller. The framework further allows for defining several properties of the eye exercise implementations, including object type, duration, speed, and background. We implemented the eye exercises using simple geometric structures and pictograms to guide the eye movements (see Figure 8 for screenshots of the pictograms). A complete list of the eye exercises is given in Table 5 and a video of the eye exercises can be found in the supplemental material.

4.2.2 Study Task. We implemented a custom VR application designed as a representative VR experience while inducing symptoms of DES. We chose a reading task, a common task used in prior work for investigating DES symptoms with digital devices [25, 86]. Because reading alone is rather uncommon in VR, we designed a word search puzzle game (*WordSearch Puzzle*) to simulate

Table 5. A List of the Eye Exercises, Including the Special Type *Closing the Eyes*

E1	Periodic Blinking	**E3**	Fixation Shift	**E5**	Smooth Pursuits	**E7**	Static Fixation	Closing the Eyes
E2	Pupillary Light Reflex	**E4**	Saccades	**E6**	Rolling Eyeballs	**E8**	Loose Focus	

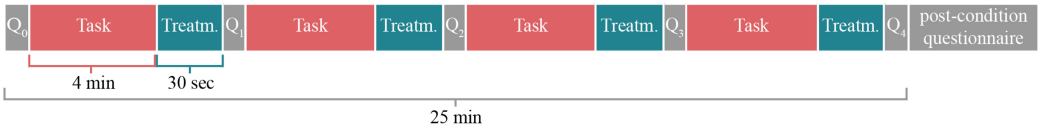


Fig. 10. Each study condition consisted of conducting the *WordSearch Puzzle* task in VR for 4 minutes, followed by a treatment method for 30 seconds. This procedure was repeated four times in each condition. In each condition, one of 10 treatment methods was active (8x eye exercises, 1x closing the eyes, and 1x control condition). We assessed 11 symptoms of DES, four statements about the eye exercises, and the number of words found at five measurement time points in each condition (Q_0–Q_4). A post-condition questionnaire was applied covering statements on the effects of the specific eye movements.

VR use, implemented as a hemisphere of letters surrounding the user (see Figure 9). This combination constitutes a trade-off between keeping the study duration relatively short and inducing eye strain, i.e., achieving measurable DES values. The users' task was to find hidden words that were distributed vertically and horizontally, and forwards and backward, in the grid. Each session hid 30 words of a pool of 240 words chosen from six categories (cities, countries, animals, means of transport, food, and common first names). Users selected words by highlighting letters with the trigger button of the controller. Correct words were highlighted in green, wrong ones in red.

4.2.3 Research Questions. We designed two user studies to evaluate the effectiveness of the eye exercises (8x eye exercises and 1x closing the eyes) to address DES in VR-HMDs. We vary the duration and frequency of eye exercises in these user studies to find the optimal application method. In the online survey, most users (60%) indicated that they would be willing to perform eye exercises for less than 1 minute in VR. When asked about the frequency, 44% indicated that they would be willing to do the exercises more than once an hour. Therefore, in the first user study, we investigate the effect of eye exercises when repeatedly applied for a short period during the exposure (high frequency, short duration). In the second user study, we investigate eye exercises when applied once during the exposure for a more extended time period (low frequency, long duration). In both studies, we compared the eye exercises to a control condition. In the control condition, participants were continuing the VR task, while in the other conditions, the main task was interrupted by breaks to perform the eye exercises. The particular research questions that we aimed at answering with these two user studies are:

RQ3: User study 2: Can DES symptoms during VR-HMD use be addressed by repeatedly applying single eye exercises during exposure for a short duration?

RQ4: User study 3: Can DES symptoms during VR-HMD use be addressed by applying a set of eye exercises once during the exposure for a longer duration?

*4.2.4 Applying Eye Exercises of Short Duration and High Frequency to Address DES in VR-HMDs (**User Study 2**).* In the first study on eye exercises, we compared the effects of the eight eye exercises and *closing the eyes* when applied of short duration and high frequency to a control condition.

Study Design. We implemented the study as a repeated-measures design with one independent variable **treatment** with 10 levels (8x eye exercise, 1x closing the eyes, and 1x control condition), resulting in 10 conditions. Each exercise was implemented to take 30 seconds. The total duration

Table 6. The Symptoms That We Measured in Both Eye Exercise User Studies

DES symptoms	burning (E), dryness (E), irritation (E), tearing, ache (I), blur, double vision, headache (I), and strain (I)
VR-specific symptoms	sensitivity to bright light, difficulty refocusing
Likert statements	My eyes feel tired/relaxed. Performing the eye movement was straining/relieving for my eyes.

(E) refers to a symptom that is externally perceived and (I) refers to a symptom that is internally perceived according to Sheedy et al. analysis [85].

of one condition was 25 minutes (Figure 10). The order of conditions was counterbalanced using a 10×10 Latin square. As we recruited 24 participants, we had imbalanced groups. We had each two participants in six of the groups, and three participants in four of the groups. Participants conducted the conditions at 10 different days to avoid carryover effects of DES symptoms across conditions. We instructed participants to perform the study application at the same time every day in order to ensure similar preconditions (e.g., screen time, general fatigue). Participants received a reward of €35 plus the chance of winning another €50 that were raffled among the five participants with the highest word count.

Measures. To measure DES subjectively, we used the nine items proposed by Sheedy et al. (burning, ache, strain, irritation, tearing, blur, double vision, dryness, and headache) [86]. In addition to these nine items, we measured two items that are specifically tailored toward DES in VR. These are "sensitivity to bright light" from Seguí et al.'s CVS-Q [28] and "difficulty refocusing" also from the CVS-Q and from the SSQ [50]. We reduced the set of items from the first user study from 21 to 11 to reduce the time that participants would need to answer the questionnaires. However, the analysis of the results could be performed on the same set of items. Similar to the blue light user study, we used a 7-point symptom severity scale to measure the symptoms. The labels of the scale were as follows: no problem (0), minimal problem (can be easily ignored without effort, (1), mild problem (can be ignored with effort, (2), moderate problem (cannot be ignored but does not influence the activity, (3), moderately severe problem (cannot be ignored and occasionally limits the activity, (4), severe problem (cannot be ignored and often limits my concentration on the activity, (5), and very severe problem (cannot be ignored and markedly limits and requires rest during the activity, (6). Lastly, we asked participants to rate the following four statements about the eye exercises on a 7-point Likert scale, reaching from *strongly disagree* to *strongly agree*: "Performing the eye exercise was straining for my eyes", "performing the eye exercise was relieving for my eyes", "my eyes feel tired", and "my eyes feel relaxed". A set of the items that were measured during each condition is shown in Table 6. All 15 items were polled five times during one condition (indicated by $Q_0 - Q_4$ in Figure 10). Once in the beginning, three times during the application, and once at the end of each condition.

After each condition participants answered a post-condition questionnaire on a laptop. This questionnaire included the following questions on a 7-point Likert scale, reaching from *strongly disagree* to *strongly agree*.

— "It was easy to perform the eye exercise."
— "The eye exercise increased my sensation of eye strain."
— "The eye exercise reduced my sensation of eye strain."
— "The eye exercise did not make a difference in my sensation of eye strain."
— "Doing the eye exercise annoyed me."
— "Doing the eye exercise was fun."
— "I enjoyed doing the eye exercise."
— "I would prefer to take off the headset to reduce eye strain rather than doing the eye exercise."
— "I would prefer to close my eyes rather than doing the eye exercise."

Furthermore, we asked them how *effectively* they perceived the eye exercise on a 7-point scale, reaching from *strong negative effect* to *strong positive effect* ("Please rate the eye exercise in terms of effectiveness from −3 to 3."). Lastly, we asked participants whether their eyes felt *more* or *less* strained after doing the eye exercise (3-point scale: *yes*, *no*, *I don't know*).

In a final questionnaire, which participants answered at the end of the study after having executed all study conditions, we asked participants to answer 13 questions about the usage of eye exercises. We asked them to choose each three eye exercises with respect to being "most effective", "least effective", "most fun", and "least fun" (e.g., "Please choose the three eye exercises that were most effective."). Furthermore, they were asked to name each three exercises that they would "most likely use" or would "most likely not use", if they were integrated into device use and if they owned a headset. Lastly, they were asked how much each of the symptoms occurred over the total duration of the study. The complete questionnaires and single item questions are given in the appendix C.1.

Participants. We recruited 24 participants (8 female, 16 male) with a mean age of 24.5 (SD = 4.3). All of them had a scientific background in being bachelor, master or doctoral students at a university. A total 10 of the participants had never used a VR-HMD before, and 13 had tried it but not used it regularly. Only one person had used a VR headset regularly before the study. All persons had normal, or correct-to-normal vision.

Apparatus and Study Application. As the participants conducted the study at home, we used mobile headsets (Oculus Go Headset). These were more easily available than stationary ones (e.g., price) and more straightforward for participants to use (e.g., standalone, no external tracking). The study task (*WordSearch Puzzle*) was implemented as a game with 10 levels. When starting the application, it showed an overview of the levels. The participants were instructed to select one specific level each day, the order of which was pre-defined by the experiment instructor. The application consisted of two parts: A word search task followed by an eye movement, which was repeated four times as shown in Figure 10.

Procedure. We introduced participants to the study in a 30-minute session, which included performing a demo level of the study application. After participants signed a consent form, we gave them an Oculus Go headset to take home for 14 days. We explained to them that they could cancel the study at any time if they felt uncomfortable without experiencing any disadvantages. We then instructed them to perform the study application once a day for 10 days. We included four additional days if they forgot to execute the application or did not feel well one day. We further gave them a list with the order of levels they had to choose. The order of levels on the list was counterbalanced among all participants. In each session, participants performed the *WordSearch Puzzle* task four times, followed by an eye exercise. In total, there were five points of measurement of eye strain in one session (see Figure 10). The questions were presented as user interfaces in VR in the same surrounding environment as the word search application and the eye movements. After participants completed the 10 days of assessment, they returned the headset, completed a final questionnaire, and received the reward.

Results. Similar to the blue light user study, we calculated four symptom means (M_{all}, M_{ex}, M_{in}, M_{vr}) to analyze the potential effects of individual eye exercises on subjectively perceived DES.[13] Furthermore, we analyzed potential effects on the four additional statements rated on a 7-point Likert scale (see Table 6 for a summary of the symptoms).

[13]We calculated relative symptom scores by subtracting the first value (Q_0) from each of the other values (Q_1–Q_4) for every symptom and participant. As such, we could eliminate potential effects that occurred due to increased entry levels.

Table 7. Results of the Non-parametric Analysis of Variances for the First User Study on Eye Exercises in Which We Investigated Eye Exercises of Short Duration and High Frequency

	M_{all}	M_{ex}	M_{in}	M_{vr}
Time	$F(1) = 23.59, p < .01, r = .49$	$F(2) = 14.63, p < .01, r = .49$	$F(1) = 24.23, p < .01, r = .49$	$F(1) = 16.15, p < .01, r = .49$
Treatment	$F(5) = 0.77, p = .58, r = .00$	$F(5) = 0.95, p = .45, r = .00$	$F(6) = 0.80, p = .57, r = .00$	$F(6) = 1.05, p = .39, r = .00$
Time:Treat.	$F(8) = 0.81, p = .59, r = .00$	$F(9) = 0.86, p = .35, r = .00$	$F(10) = 1.11, p = .18, r = .27$	$F(7) = 0.83, p = .56, r = .00$
	$M_{straining}$	$M_{relieving}$	M_{tiring}	$M_{relaxing}$
Time	$F(2) = 6.99, p < .01, r = .49$	$F(2) = 2.83, p = .07, r = .28$	$F(2) = 16.80, p < .01, r = .49$	$F(1) = 17.23, p < .01, r = .49$
Treatment	$F(5) = 5.23, p < .01, r = .57$	$F(4) = 5.47, p < .01, r = .57$	$F(6) = 0.49, p = .81, r = .00$	$F(5) = 1.2, p = .31, r = .00$
Time:Treat.	$F(8) = 1.2, p = .29, r = .00$	$F(9) = 0.82, p = .60, r = .00$	$F(10) = 1.0, p = .44, r = .00$	$F(8) = 1.11, p = .35, r = .00$

The data were analyzed following a two-factorial design with one factor **treatment** and one factor **time**. For the four symptom means M_{all}, M_{ex}, M_{in}, and M_{vr} we analyzed the relative symptom scores. Therefore, the factor **time** had four levels Q_{1-0}, Q_{2-0}, Q_{3-0}, and Q_{4-0}. For the four additional statements $M_{straining}$, $M_{relieving}$, M_{tiring}, and $M_{relaxing}$, we analyzed the absolute values with the factor **time** having four levels Q_1, Q_2, Q_3, and Q_4, as we excluded the first value at Q_0 from the analysis. We report r-equivalent effect sizes based on Rosenthal and Rubin's method [82].

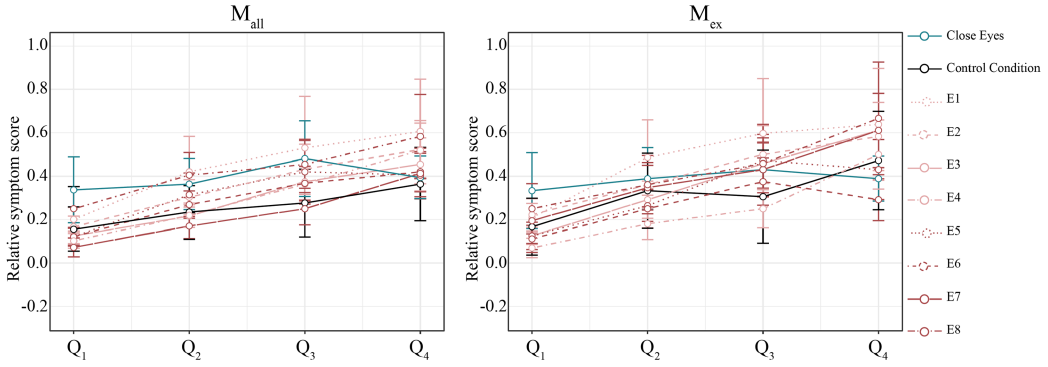
We found good values for the internal consistency (Cronbach's alpha) for all four symptom means: $M_{all} : \alpha = 0.92$, $M_{ex} : \alpha = 0.84$, $M_{in} : \alpha = 0.82$, and $M_{vr} : \alpha = 0.81$. As some of the groups significantly differed from a normal distribution (refer to the appendix Section C.2.1 for normality test results), we analyzed potential effects using a two-factor non-parametric test for variance analysis with the factors **treatment** (10 levels: 8x eye exercises, 1x closing the eyes, and 1x control condition) and **time** (four levels: Q_{1-0}–Q_{4-0}). To calculate the test, we used the nparLD R-package [67]. In the following, we report ANOVA-type statistics and effect sizes calculated according to Rosenthal and Rubin's method [82]. For post-hoc tests, we applied Dunn-tests with Bonferroni correction.

We found significant effects of **time** on all four means ($M_{all}, M_{ex}, M_{in}, M_{vr}$), indicating that the symptoms became more severe the longer the exposure was (see Table 7 for test statistics and Figure 11 for the course of the symptom means). We did not find a statistically significant effect of **treatment**, nor an interaction effect between **time** and **treatment**, on any of the four relative symptom means. In summary, we found that the severity of the symptoms increased over time in all conditions, but the treatment did not have an effect on the experienced symptom level.

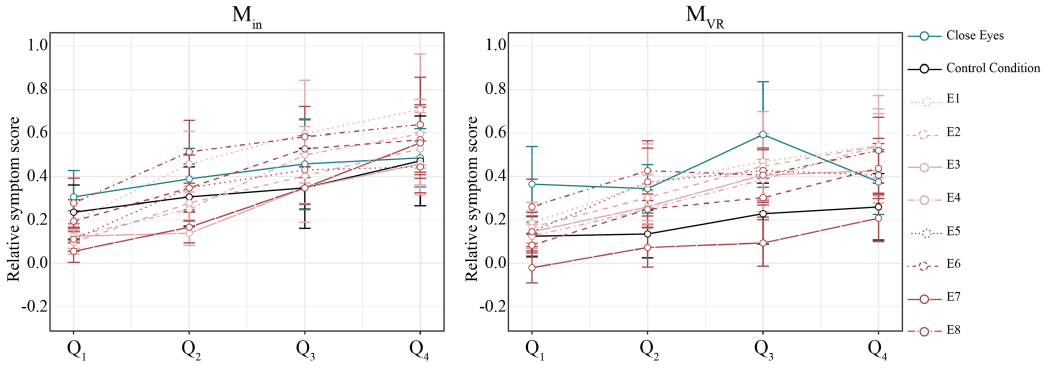
For the perceived feeling of strain and relief,[14] we found a significant effect of **treatment**, as well as an interaction effect between **time** and **treatment** (see Table 7 for test statistics). For the perceived tiring or relaxing effect of the exercises, we found a statistically significant effect of **time**, but not **treatment**. Lastly, for the presence item we found a statistically significant effect of **time** ($F(2) = 6.59, p < .01, r_e = 0.98$), showing that presence scores increased with exposure time from $M_{Q0} = 3.33$ ($SD = 0$) to $M_{Q4} = 3.83$ ($SD = 1.42$).

In the following, we report the results for the post-condition questionnaire. The participants agreed that all eye exercises, including *closing the eyes*, were easy to perform with the highest values for closing the eyes ($M = 5.8, SD = 1.7$) and **E8** "loose focus" ($M = 5.6, SD = 1.5$), and the lowest value for **E6** "rolling the eyeballs" ($M = 4.2, SD = 1.7$). The majority of participants stated that their eyes felt more strained for all eye exercises after the VR task than before the VR task ($M = 15.78, SD = 2.25$). There was also great agreement that their eyes did not feel less strained than before the VR experience ($M = 18.7, SD = 2.16$). When looking at the general questions that participants answered after each condition, participants are split about the effects of coping strategies. When asked whether doing the eye exercises annoyed them, E6 (13 participants rated this statement with at least *slightly agree*) and E7 (12 participants rated this statement with at

[14]For these values, we did not calculate relative changes, but used absolute values. We excluded the first value, measured at Q_0, from the analysis because participants did not perform an eye exercise before the first task.



(a) We show the relative symptom score for the overall symptom mean M_{all} and the external symptom mean M_{ex}.



(b) We show the relative symptom score for the internal symptom mean M_{in} and the VR-specific symptom mean M_{vr}.

Fig. 11. Relative symptom scores of the four symptom means that were assessed during each condition in the first user study on eye exercises. The values depict relative changes to the first measurement time point Q_0. Bars represent standard deviation.

least *slightly agree*) were the ones that received the most negative ratings. For all other conditions, the answers did not indicate a clear negative aspect. When asked whether participants would enjoy doing the eye exercises, on average eight rated the statement with *slightly agree* or higher, nine rated the statement with *slightly disagree* or lower,and seven rated it as *neutral*. Opinions on whether participants preferred to take off the headset than doing the eye exercises were also mixed. On average 12 indicated that they would prefer to take off the headset, eight indicated that they would prefer to do the eye exercises, and four rated the statement as neutral. All results of the post-condition questionnaires are given in the appendix Section C.2.2.

In the final questionnaire, we asked participants to choose each three eye exercises concerning "effectiveness" and whether they would use them to reduce eye strain during a VR experience (e.g., "Please choose the three eye exercises that were most effective."). *Close eyes* was named 22 times to be the most effective eye exercise in reducing eye strain during the experience, followed by *periodic blinking* (9) and *none* (9). When asked for the three least effective eye exercises, *none* was named 13 times, followed by *pupillary light reflex* (11) and *rolling the eyeball* (10). When asked which eye exercises participants would most likely use, *close eyes* was named 18 times, followed by *none* (15) and *smooth pursuit* (9). The ones that participants would most likely not use were *none* (13), *rolling the eyeball* (12), and *pupillary light reflex* (10). Lastly, we asked participants to rate the occurrence of single symptoms over the total duration of the study. Two-thirds of the participants stated that

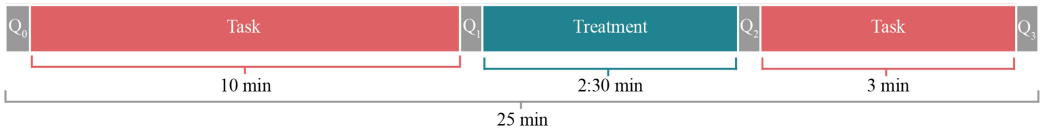


Fig. 12. The second user study on eye exercises consisted of three conditions. Participants first played the *WordSearch Puzzle* for 10 minutes as the task for each condition. They then performed one of three treatment methods for 2:30 minutes. After the treatment, participants completed the second *WordSeach Puzzle* task of three minutes. The three treatment methods were *eye exercises*, *closing the eyes*, and the *control condition* in which participants continued with the task. We recorded objective and subjective dependent variables at four measurement time points (Q_0–Q_3).

symptoms occurred up to 25% of the time, with the following ones being experiences particularly often: *strain*, *dryness*, *blurred vision*, and *sensitivity to bright light*.

*4.2.5 Applying Eye Exercises of Long Duration and Low Frequency to Address DES in VR-HMDs (**User study 3**).* As we did not find statistically significant effects of eye exercises of short duration and high frequency on DES symptoms, in a second step, we tested their effectiveness when applied of long duration and low frequency. Instead of applying one eye exercise for a longer duration, we put together five individual eye exercises to a set. The set of eye exercises constitutes a trade-off between a longer duration and not performing the same eye exercise for a prolonged time period. By putting together a set of five exercises a 30 seconds, we achieved a total treatment time of 2:30 minutes. We selected the eye exercises **E1** periodic blinking, **E2** pupillary light reflex, **E3** fixation shift, **E4** saccades, and **E6** rolling eyeball into one direction. Again, we considered *closing the eyes* a special case of eye exercise, as it does not provide a visual stimulus. Therefore, we added it as a separate condition to the study.

In summary, in the second user study on eye exercises, we investigated two treatment strategies: *closing the eyes* and a *set of eye exercises* in comparison to a control condition. We also added objective eye measures (pupil size and blink rate) to the study design to gain a more detailed impression of DES symptoms.

Study Design. The study was implemented as a repeated-measures design with one independent variable **treatment** with three levels (*eye exercises*, *closing the eyes*, and *control condition*) resulting in three conditions. The condition order was counterbalanced using a 3x3 Latin square. As we recruited 19 participants, the grouped were imbalanced. We had two groups with each six participants, and one group with seven participants. Participants completed the three conditions at three different days to ensure that there were no carry over effects of DES. As dependent variables we assessed objective (pupil size and blink rate) and subjective measures of DES (11 symptoms on a 7-point scale). An overview of the study design is shown in Figure 12. We used the same task as in the previous user study to simulate VR usage and induce DES symptoms (see Figure 9 for a screenshot of the *WordSearch Puzzle*).

We split each condition into three parts to measure the immediate and extended effects of the treatment method. The procedure for each condition of the study is shown in Figure 12. First, participants conducted a round of the *WordSearch Puzzle* game as the initial task for 10 minutes. Afterward, in each condition, they performed one of the three treatment methods *eye exercises*, *closing the eyes*, and *control condition* for 2:30 minutes. In the third part, to induce DES symptoms a second time, participants completed a second round of the *WordSearch puzzle* for 3 minutes. One condition took about 30–40 minutes and participants received a reward of € 25 in total. For an

additional incentive to take part in a longitudinal study over three days, we raffled another € 50 among the five participants with the highest word score after all three sessions.

Measures. Similar to the two previous user studies, we measured 11 DES symptoms on a 7-point scale, reaching from *no discomfort* to *very severe discomfort*. In addition to the nine items of Sheedy et al. [86], we assessed the two additional VR-specific items "sensitivity to bright light" and "difficulty refocusing". Lastly, participants were asked how tired and relaxed their eye felt after each condition, on a 7-point scale.

As dependent objective variables, we used eye tracking measures. We chose this method since eye trackers will most probably be integrated into future HMDs [40] and, as such, enable users to measure DES objectively without requiring additional instrumentation. Wang et al. suggested that DES can be detected based on blinking metrics [97]. Additionally, we recorded pupil size as an indicator for general fatigue [98]. The Pupil Capture software allows for adjusting confidence and filter length values for blinks to be detected. To identify correct values, we followed the procedure of Langbehn et al. [60]. Similarly, we conducted a pre-test with three persons identifying that a confidence value of 0.5 (ranging from 0 to 1 with 0 meaning no confidence and 1 representing the highest possible confidence) and a duration filter length of 300 ms would give the best result for recognizing blinks. Consequently, we recorded *blinks* and *pupil size* of both eyes of the participants as objective dependent variables.

Participants. We recruited 19 participants (7 female, 12 male) with an average age of 26.9 (SD = 2.5). They were students (13) or employed for wages (6). Only three participants had never used VR before, the others either currently used (2) or had used a headset regularly (5). The remaining 9 participants had tried VR before, but never used it regularly.

Apparatus and Study Application. The study was implemented using Unity 3D and the HTC Vive Pro with integrated Pupil Labs Add-ons [49]. We used version 1.6.11 of the Pupil Capture and Player Software. We used the same study application as in the previous user study on eye exercises. However, this time, the eye exercises were not presented individually, but in a set. A video of the three conditions is given in the supplemental material.

Procedure. The study took place in a quiet room at a university. After a brief introduction, participants signed a consent form and went through a training session of the *WordSearch Puzzle* study task. After completing eye tracker calibration, users underwent a 20-second assessment of objective measures (Q_0 in Figure 12).

During measurement, participants looked at a fixation cross with a dark background instructed to internally count up from one to control cognitive load and affective factors that influence pupil size [20, 70]. After measurement, participants answered the aforementioned 13 items on 7-point Likert scales while staying inside VR. These were implemented as user interfaces in VR in the same virtual surroundings as during the task. After that, participants performed the *WordSearch Puzzle* task for 10 minutes (see Figure 12), followed by a second assessment of DES (Q_1). Then, they were presented with one of the three treatment variants per condition (see "Treatment" in Figure 12), where each took 2:30 minutes. After the third assessment of DES (Q_2), participants had to complete another round of the *WordSearch Puzzle* for 3 minutes, followed by the fourth assessment of DES (Q_3).

Results. The dependent measures were gathered at four time points Q_0–Q_3 as depicted in Figure 12.

Similar to the first user study on eye exercises, we calculated relative symptom scores of the subjective data by subtracting the first value (Q_0) from the three values at Q_1–Q_3.

Table 8. Results of the Non-parametric Analysis of Variances for the Second User Study on Eye Exercises in Which We Investigated Eye Exercises of Long Duration and Low Frequency

	M_{all}	M_{ex}	M_{in}	M_{vr}
Time	$F(2) = 13.65, p < .01, r = .54$	$F(2) = 5.88, p < .01, r = .54$	$F(2) = 17.68, p < .01, r = .54$	$F(2) = 9.91, p < .01, r = .54$
Treatment	$F(2) = 3.41, p = .04, r = .40$	$F(2) = 3.18, p = .04, r = .40$	$F(2) = 1.82, p = .16, r = .21$	$F(2) = 2.25, p = .11, r = .26$
Time:Treat.	$F(3) = 4.87, p < .01, r = .59$	$F(3) = 3.86, p < 0.01, r = .59$	$F(2) = 4.35, p < .01, r = .59$	$F(3) = 3.20, p = .02, r = .51$

We report ANOVA-type statistics. The data were analyzed following a two-factorial design with one factor **treatment** and one factor **time**. We analyzed the relative symptom means. Therefore, the factor **time** has three levels. We report r-equivalent effect sizes.

Table 9. Mean Values for the Four Symptom Means of the Second User Study on Eye Exercise, Summarized by Condition

	M_{all}	M_{ex}	M_{in}	M_{vr}
Close Eyes	$M = 0.84, SD = 0.85$	$M = 1.25, SD = 1.35$	$M = 1.11, SD = 1.15$	$M = 0.30, SD = 0.65$
Eye Exercises	$M = 0.91, SD = 0.82$	$M = 1.35, SD = 1.20$	$M = 1.09, SD = 0.88$	$M = 0.38, SD = 0.83$
Control Condition	$M = 1.18, SD = 0.87$	$M = 1.68, SD = 1.15$	$M = 1.29, SD = 0.87$	$M = 0.83, SD = 1.03$

We then calculated the same four symptom means as in the two previous user studies (M_{all}, M_{ex}, M_{in}, M_{vr}) to analyze the potential effects of the treatment methods on subjectively perceived DES. We found acceptable internal consistency values (Cronbach's alpha) for M_{all} ($\alpha = 0.88$), M_{ex} ($\alpha = 0.76$), and M_{vr} ($\alpha = 0.74$). Only the value for M_{in} was slightly below the threshold to indicate acceptable internal consistency ($\alpha = 0.69 < 0.7$).

We analyzed the data with a two-factorial test with the factors **treatment** (three levels: eye exercises, closing the eyes, and control condition) and **time** (three levels: Q_{1-0}, Q_{2-0}, Q_{3-0}).[15] As some of the variables differed from a normal distribution, we tested the symptom means with a non-parametric test for analysis of variances (see Shapiro–Wilk normality test results in the appendix Section D.1). We found a significant effect of **treatment** on the overall symptom mean M_{all} and the external symptom mean M_{ex} (see Table 8 for test statistics). Post-hoc tests for the overall symptoms mean M_{all} showed that the *control condition* produced significantly more severe DES symptoms than the *closing the eyes* condition (see Table 9 for the mean values of the four symptom means in each condition). For the external symptoms mean M_{ex}, we could not observe any significant group differences. Furthermore, we found an effect of **time** on all four symptom means. Post-hoc tests only revealed one group difference, showing that the internal symptom mean value M_{in} was significantly lower for the time point Q_2 than for the time point Q_3. Lastly, we found interaction effects between **time** and **treatment** for all four symptom means. All results for the non-parametric analysis of variances are given in Table 8. A time course of the four symptom means averaged across all participants is shown in Figure 13.

The interaction effects between **time** and **treatment** for all four symptom means suggest that there is an effect of the different treatment conditions depending on the measurement time point. To reveal such potential time-dependent effects, we followed up with a more detailed analysis of the symptom means. In particular, we analyzed potential differences in symptom means directly *after* the treatment (Q_2) versus directly *before* the treatment (Q_1). This would reveal a potential immediate effect of the treatment method. To do so, we tested whether **treatment** as independent variable had an effect on the relative symptom means of **Dif(Q_2-Q_1)** as dependent variables. Secondly, we analyzed potential extended **treatment** effects by building an extended symptom mean

[15]We removed Q_0 from the analysis, as we calculated relative symptom scores based on this value.

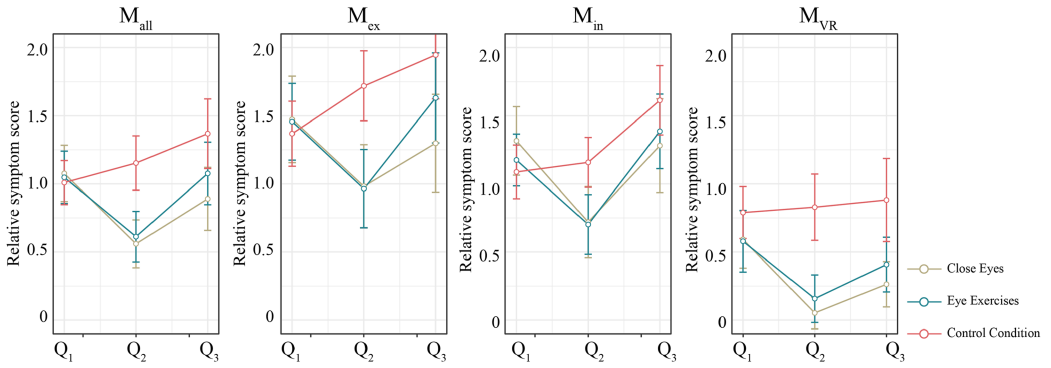


Fig. 13. The values for the four relative symptom mean over time from Q_{1-0} to Q_{3-0} averaged across all participants for user study 3. Colors indicate the different treatment methods *eye exercises*, *closing the eyes*, and the *control condition*. M_{all} is the mean of all 11 symptoms. M_{ex} is the mean of the external symptoms burning, dry eye, and irritation. M_{in} is the mean of the internal symptoms eye ache, headache, and eye strain. M_{vr} is the mean of the VR-specific symptoms blurred vision, difficulty focusing, double vision, and sensitivity to bright light. Bars represent standard deviation.

Table 10. Mean Values for the Four Relative Symptom Means of the Second User Study on Eye Exercises, Calculated as Indicating an Immediate Effect **Dif**(Q_2-Q_1) and an Extended Effect **Dif**(Q_3-Q_1)

	M_{all}	M_{ex}	M_{in}	M_{vr}
Close Eyes (Q_2-Q_1)	$M = -0.52, SD = 0.56$	$M = -0.49, SD = 0.78$	$M = -0.60, SD = 0.86$	$M = -0.54, SD = 0.77$
Eye Exercises (Q_2-Q_1)	$M = -0.44, SD = 0.60$	$M = -0.49, SD = 0.96$	$M = -0.47, SD = 0.82$	$M = -0.42, SD = 0.59$
Control Condition (Q_2-Q_1)	$M = 0.14, SD = 0.38$	$M = 0.35, SD = 0.50$	$M = 0.07, SD = 0.48$	$M = 0.04, SD = 0.56$
Close Eyes (Q_3-Q_1)	$M = -0.19, SD = 0.44$	$M = -0.18, SD = 0.86$	$M = -0.04, SD = 0.78$	$M = -0.33, SD = 0.58$
Eye Exercises (Q_3-Q_1)	$M = 0.03, SD = 0.56$	$M = -0.18, SD = 0.87$	$M = 0.21, SD = 0.88$	$M = -0.17, SD = 0.43$
Control Condition (Q_3-Q_1)	$M = 0.36, SD = 0.62$	$M = 0.58, SD = 0.71$	$M = 0.53, SD = 0.73$	$M = 0.09, SD = 0.85$

of the time points after the second task (Q_3) and before the first task (Q_1). All mean values of these variables are shown in Table 10.

We first report results for the *immediate* effect of the treatment method on the symptoms means (Q_2-Q_1). Boxplots of these results are presented in Figure 14. Using a one-way repeated-measures ANOVA, we found a significant effect of **treatment** on the mean of all symptoms M_{all} ($F(2, 36) = 7.79, p < .01, \eta^2 = .25$) and on the external symptom mean M_{ex} ($F(2, 36) = 6.24, p < .01, \eta^2 = 0.22$). Post-hoc tests revealed that *close eyes* and *eye exercises* resulted in less severe overall symptom means M_{all} in comparison to the *control condition* (close eyes: $p < .01$, eye exercises: $p = .028$) and in less severe external symptom means M_{ex} in comparison to the *control condition* (close eyes: $p < .01$, eye exercises: $p = .01$). The internal and VR-specific symptoms means were analyzed with Friedman tests. As effect size we report Kendall's W. For the internal symptoms mean M_{in}, we found a statistically significant small effect of **treatment** method ($\chi^2(2) = 9.79, p < .01, W = 0.14$). Pairwise Wilcoxon signed rank test with Bonferroni correction between groups revealed a statistically significant difference between the *close eyes* and the *control condition* ($p = .01$). For the VR-specific symptom mean M_{vr}, we did not find a statistically significant effect of treatment method ($\chi^2(2) = 5.26, p = .07$).

In the following, we report results of the *extended* effect of the treatment methods on the symptom means. Boxplots of the results are shown in Figure 15. The overall symptom mean M_{all}, internal symptom mean M_{in}, and VR-specific symptom mean M_{vr} were tested with Friedman tests. We found statistically significant small effects of **treatment** method on the overall symptom mean

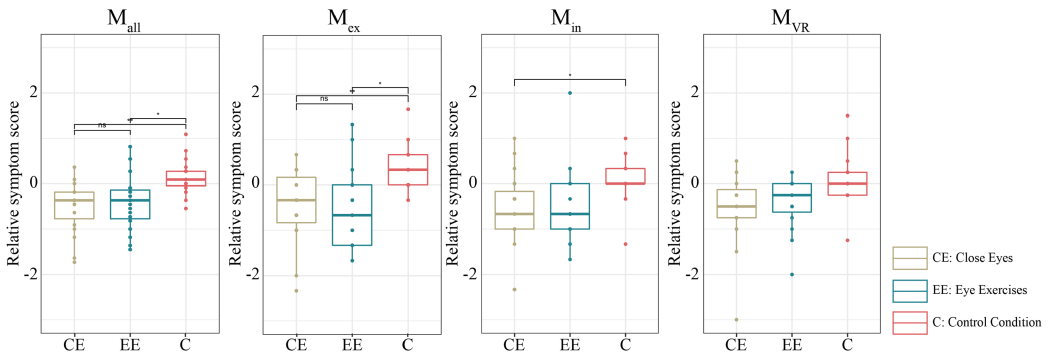


Fig. 14. Boxplots of the relative values of the four symptom means, averaged over all participant and presented separately for each treatment method. The values reflect the *immediate* effect of treatment method on DES symptoms (**Dif(**Q_2-Q_1**)**).

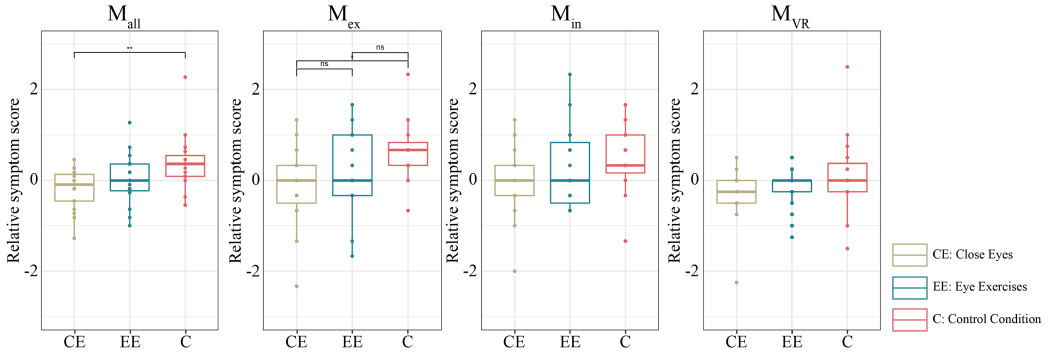


Fig. 15. Boxplots of the relative values of the four symptom means, averaged over all participant and presented separately for each treatment method. The values reflect the *extended* effect of treatment method on DES symptoms (**Dif(**Q_3-Q_1**)**).

M_{all} ($\chi^2(2) = 9.95$, $p < .01$, $W = 0.26$) and the internal symptom mean M_{in} ($\chi^2(2) = 10.0$, $p < .01$, $W = 0.26$), but not on the VR-specific symptom mean M_{vr} ($\chi^2(2) = 1.3$, $p = .52$). Post-hoc tests revealed a statistically significant difference between the overall symptom mean M_{all} of the *close eyes* and the *control condition* ($p < .01$). We did not find any other group differences for the overall symptom mean M_{all} or the internal symptom mean M_{in}. The external symptom mean M_{ex} was tested with a one-way repeated-measures ANOVA for the potential effects of the treatment method. We found statistically significant differences between the **treatment** methods ($F(2, 36) = 4.75$, $p = .02$, $\eta^2 = 0.13$). Post-hoc analyzes with a Bonferroni adjustment revealed a statistically significant difference between *closing the eyes* and the *control condition* ($p = .02$).

In summary, we found that the eye exercise and closing the eyes treatments reduced DES symptoms compared to the control condition, particularly shown for the symptom means M_{all} and M_{ex}. We were able to show this for the immediate and the extended measure time points.

The analysis of the objective measures will also focus on potential immediate (**Dif(**Q_2-Q_1) and extended (**Dif(**Q_3-Q_1) effect of the treatment method. We recorded pupil size and blink data of both eyes of the participants. In each frame, the eye tracker assigns a confidence value $c \in [0 : 1]$

to the data to indicate the quality of detection. Pupil Labs recommends to use data with $c \geq 0.6$.[16] However, we chose a value of $c \geq 0.8$ as an indicator of good quality data based on previous experience with the eye tracker. We had to exclude several time segments (>5 seconds) per participant from evaluation due to confidence values being constantly below 0.8. For the statistical analysis of the immediate **Dif(Q_2−Q_1)** and the extended effect **Dif(Q_3−Q_1)**, we only included participants that had no missing time segments in Q_1, Q_2, and Q_3, which resulted in a number of 11 participants. For the analysis of pupil size values, we cut off the first 5 seconds from the 20-second segments at each time point to exclude the pupil's adaptation phase to changing lighting conditions. We interpolated missing values by calculating the mean of the two values before and after a missing value.

We followed Klingner's approach of baseline subtraction for pupil size normalization [54]. We took pupil size values at Q_1 as the baseline, as we were interested in whether pupil size values at Q_2 and Q_3 changed in relation to those directly after the task. Pupil size during the segments Q_2 and Q_3 was then normalized by subtracting the baseline mean from each value. The values were then averaged over the whole time frame of 15 seconds per participant and time segment. We compared baseline-corrected pupil size values of the different conditions using a repeated-measures ANOVA with one factor (treatment). The tests we report in the following were conducted for the participants' left eyes since we had better confidence values for the left than for the right eye.

For the immediate effect of the exercises (**Dif(Q_2−Q_1)**), we found an effect of **treatment** on pupil size changes ($F(2, 28) = 14.23$, $p < .001$, $\eta^2 = 0.5$). A post-hoc t-test with Bonferroni correction revealed that pupil size was significantly smaller for ***close eyes (CE)*** than for ***eye exercises (EE)*** ($M_{CE} = -6.02$, $SD_{CE} = 1.84$, $M_{EE} = 1.1$, $SD_{EE} = 1.98$, $p < .001$) and the ***control condition (CC)*** ($M_{CC} = 1.13$, $SD_{CC} = 1.12$, $p = .003$). We did not find an extended effect (**Dif(Q_3−Q_1)**) of the treatment methods on pupil size.

For blinking data, we chose a different data processing procedure than for pupil size, as these values depend significantly more on the confidence values of the tracker. A low confidence value is interpreted as a closed eye. As such, a blink is detected when the confidence value drops below the threshold for the duration of the filter length. An interpolation for missing values is complicated. For a highly fluctuating confidence value, it is difficult to differentiate whether the data quality of the trial was poor or the person blinked a lot. Therefore, we decided to manually label the blinking data by applying a binary decision (blink yes/no). A Friedman test did not reveal an effect of treatment on blinking data for an immediate (**Dif(Q_2−Q_1)**) or extended effect (**Dif(Q_3−Q_1)**). We further analyzed a possible decrease in blink rate during the study, which we did not find.

4.2.6 Discussion. We conducted two user studies to evaluate the effectiveness of eye exercises and to determine the frequency and duration of this potential treatment method. In the first user study on eye exercises, we analyzed single eye exercises when applied with high frequency and low duration. In the second user study on eye exercises, we analyzed the effect of a set of five eye exercises when applied only once (low frequency) and for a long duration. Similar to previous studies [58, 71, 84] and the blue light study, we found that DES symptoms increased significantly over 25 minutes in both studies, which confirms that users already experience DES after a short time of usage. Furthermore, we found that external and internal symptoms increased significantly in both studies. In the second user study, external symptoms increased more severely than internal ones, which stands in contrast to studies suggesting that VR-HMDs dominantly cause internal DES symptoms [103]. These observations could be attributed to the VR task (reading), which probably

[16]Pupil Labs: https://docs.pupil-labs.com/core/software/pupil-player, last retrieved: April 27, 2021.

increased irritation of the ocular surface and, therefore, caused an increase in external symptoms. The results further show that symptoms have occurred consistently over 10 (user study 2) and three (user study 3) days.

On Eye Exercises of Short Duration and High Frequency. In the first user study, we could not find effects between different eye exercises when assessed with the symptom severity scale. However, when directly asking the participants whether the eye exercises felt *straining* or *relieving* directly after each condition, we found that *closing the eyes* had a statistically significant positive effect on participants' eyes. As we did not find such an effect with the symptom severity scale and the items that were assessed during the application, we discuss two potential interpretations. On one hand, the effect could be justified by the different rating scales. The post-exposure questions were assessed on a 7-point Likert scale, reaching from *strongly disagree* to *strongly agree*, while the questions that were assessed during the experiment were assessed on a 7-point symptom severity scale, reaching from *no problem* to *very severe problem*. The different results indicate that the Likert scale is either easier for participants to assess or is more sensitive to detect an effect. This would be somewhat surprising, as discomfort is usually assessed with unipolar scales [39], similar to the symptom severity scale that we used. Likert scales are bipolar scales, and our results indicate that they might be better to assess symptoms of discomfort with regards to the users' eyes. However, we also have to consider that in the post-experiment questionnaire, we did not ask for the rating of symptoms but the perceived effect of coping strategies. Therefore, such a rating scale can only be applied to these specific cases. Another interpretation of the effect could be that the Likert scale was more sensitive in detecting an effect in comparison to the symptoms severity scale. This is, however, unlikely, as the severity scale presents seven items of one direction, and the Likert scale presents only three items for each direction, which could indicate that seven points are too little to assess an effect in symptomatology.

Another explanation for why we could not detect a difference between the eye exercises could be that participants did not correctly perform the eye exercises. However, as the vast majority agreed that all of the exercises were easy to perform and that they even had fun doing the exercises, this is rather unlikely (see "It was easy to perform the eye exercise"). In the subjective ratings about the eye exercises, participants' answers were somewhat mixed. Some had strong positive or negative opinions about the effect of the strategies, while others rated them as neutral. We could observe these differences in evaluation for all of the eye exercises, even for **E4** (saccades) and **E7** (static fixation)—although it is known that large saccades and prolonged fixation duration can cause eye strain [6]. When asked whether participants would prefer to close their eyes over doing the exercise, similarly, about half of the participants agreed to the statement, and the other half disagreed. Interestingly, in particular, for this point, there were very few neutral answers. In summary, these subjective results suggest that participants not only experience DES differently but also perceive eye exercises to alleviate it in different ways. This might indicate the existence of different susceptibility groups. If these differences in susceptibility exist, this could be an explanation for why the eye exercises were rated in different ways, as they were only beneficial for participants who experienced symptoms.

On Eye Exercises of Long Duration and Low Frequency. In contrast to the first user study on eye exercises, in the second user study, we found a positive effect of both *eye exercises* and *closing the eyes* on the perceived severity of DES symptoms in comparison to the control condition. We found that eye exercises effectively reduce overall and specifically external symptoms when measured immediately after their application. The positive effect of eye exercises on external symptoms even persists after a second straining VR task is applied. Closing the eyes affects overall symptoms, external and internal symptoms positively when measured immediately after application. It also

has an extended effect on the overall and external symptoms of DES. In summary, when applied for 2:30 minutes directly after the straining VR experience, eye exercises effectively reduce overall DES symptoms. Furthermore, both eye exercises and closing the eyes seem to be effective for external symptoms related to the irritations of the ocular surface of the eye (burning, irritation, and dryness). This is even though only one of five eye exercises (periodic blinking) was specifically designed to reduce external symptoms. Therefore, we assume that the proposed eye exercises cannot be assigned exclusively to one category of symptoms each, but rather that they have a more general effect on participants' experience of DES.

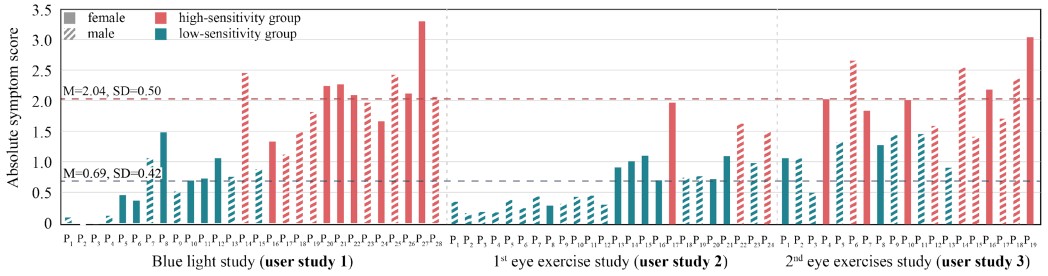
While we could observe effects for the subjective results, for the objective measures, we only observed a decrease in pupil size from baseline level for the close eyes condition in comparison to eye exercises and the control condition. A decrease in pupil size was reported as an indicator of eye strain [42]. However, we found a reduction of symptoms after closing the eyes for all symptom factors. Therefore, we assume that a lower cognitive load after closing the eyes than the eye exercises and the word search task [92] or a prolonged phase of pupil size adaption to the background after opening the eyes [100] is the reason. In comparison to the first user study on eye exercises, we conclude that eye exercises and closing the eyes effectively reduce DES when being applied for a longer time period. This speaks for integration of eye exercises (including closing the eyes) into VR experiences either by integrating the mechanism constantly into user interfaces or by integrating them as breaks, e.g., during loading screens or similar.

On the Internal Consistency of the Symptom Means. We summarized the single symptoms to four symptom means for all three user studies and analyzed the internal consistency of the mean values with Cronbach's alpha. In general, we found acceptable (>0.7) to good (>0.8) internal consistency values for the four symptom means in all three studies, indicating that the symptoms that we grouped as *external*, *internal*, or *VR-specific* show a meaningful correlation. Overall, the internal consistency was lower for internal symptoms than for external ones, which coincides with Sheedy et al.'s findings on weaker correlations for their internal symptoms [86]. We had only one poor value of 0.51 for the VR-specific symptom mean in the blue light study. However, for the other two user studies, the α-value for the VR-specific was acceptable and good. This shows that the symptoms *blurred vision*, *difficulty focusing*, *double vision*, and *sensitivity to bright light* reflect well on a common underlying construct. We labeled this construct as VR-specific in our studies. However, as we have not investigated specific inducing conditions and related them to the symptom means, we can only hypothesize that these symptoms reflect on a VR-specific factor. Overall the α-values were lowest for the blue light study and highest for the study on eye exercises that participants conducted at their homes. The generally lower values of internal consistency for the blue light study could indicate that participants who are recruited using an online platform might be less careful in selecting the adequate value on the rating scales.

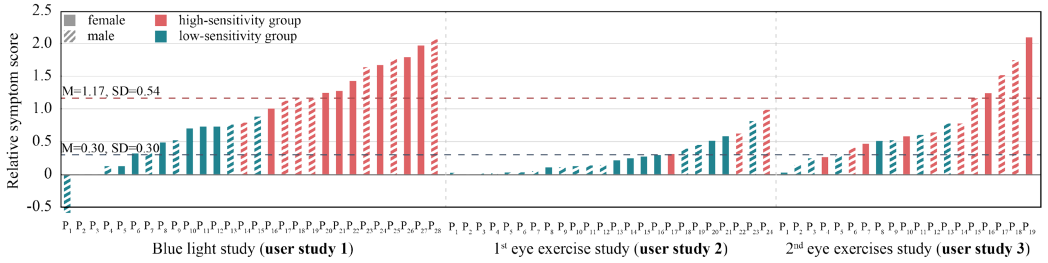
5 ANALYSING DES IN VR HMD USERS

5.1 Analysis of Influence Factors on DES in VR-HMDs

We found a high variability of DES scores among participants in all three user studies presented in this article. Some participants seemed not to experience even slight symptoms even after continuous use, while others experienced severe symptoms already after a short usage duration. The online survey results confirm these findings: Some respondents mentioned not being affected, while others experienced severe symptoms very frequently when they used their device. These observations indicate differences in susceptibility to DES, i.e., the existence of groups of users, some of which experience symptoms stronger than others. Previous work made similar observations while studying simulator sickness, which is related to DES as another form of discomfort in VR [39].



(a) One absolute mean value, calculated for all participants of the three user studies.



(b) One mean value, calculated for all participants of the three user studies.

Fig. 16. We calculated one absolute and one relative mean value for all participants of the three user studies, sorted by user study and relative symptom mean. For the blue light study, the mean value was calculated as the mean value of all measurement time points averaged over all three conditions. For the two user studies on eye exercises, we calculated a mean value of the 11 DES symptoms averaged over all conditions and measurement time points.

Several studies have confirmed that different people experience simulator sickness to different extents [84]. These differences have often been attributed to sex as one decisive influence factor [33], and several studies found that women seem to experience simulator sickness more severely than men [3, 8, 34]. However, others suggest that these differences in sickness should be attributed to susceptibility rather than sex [21]. We performed additional analyzes on our user study data to gain more clarity on both potential influence factors on DES.

5.1.1 Data Preparation. First, we calculated two mean DES scores for each participant of all three user studies: one absolute DES score (M_a) and one relative DES score (M_r). We only considered the within-experiment measures, as we did not gather post-experiment measures in the second eye exercise user study. For the blue light study, we calculated M_a as the mean of all single-item DES measurements Q_0–Q_{10}, while M_r is the mean of the relative DES scores per participant $Q_1 - Q_0$–$Q_{10} - Q_0$. To conduct the process equally across all three user studies, we chose to use always the first value (Q_0) as the baseline. For the first user study on eye exercises, we calculated M_a as the mean value of all 11 single symptoms, averaged across all measurement time points Q_0–Q_4. For M_r, we first calculated a relative values separately for each of the 11 symptoms. Then the mean values of all symptoms at Q_0 were subtracted from the measurements at Q_1–Q_4, separately for each participant. For the second user study on eye exercises, M_a was calculated similarly to the first user study on eye exercises, i.e., the 11 symptoms were averaged to one absolute mean value for each participant. We calculated M_r by first subtracting the value at Q_0 from every value Q_1–Q_3 for each of the 11 symptoms. Afterward, a mean of all relative symptom means was calculated. The mean values for each participant are presented in Figure 16.

Table 11. Summary Statistics of the Absolute and Relative Symptom Means for All Participants of All Three User Studies

Susceptibility	Sex	Experiment	N	Variable	Mean	SD	Variable	Mean	SD
high	f	blue light study	7	M_a	2.15	0.61	M_r	1.48	0.34
high	f	1^{st} eye exercise study	1	M_a	1.99	N/A	M_r	0.31	N/A
high	f	2^{nd} eye exercise study	5	M_a	2.24	0.48	M_r	0.92	0.75
low	f	blue light study	7	M_a	0.68	0.48	M_r	0.45	0.30
low	f	1^{st} eye exercise study	7	M_a	0.85	0.29	M_r	0.31	0.17
low	f	2^{nd} eye exercise study	2	M_a	1.19	0.15	M_r	0.27	0.34
high	m	blue light study	7	M_a	1.90	0.48	M_r	1.38	0.45
high	m	1^{st} eye exercise study	2	M_a	1.58	0.10	M_r	0.80	0.26
high	m	2^{nd} eye exercise study	6	M_a	2.06	0.54	M_r	1.04	0.52
low	m	blue light study	7	M_a	0.49	0.43	M_r	0.30	0.50
low	m	1^{st} eye exercise study	14	M_a	0.44	0.25	M_r	0.16	0.24
low	m	2^{nd} eye exercise study	6	M_a	1.13	0.37	M_r	0.42	0.24

5.1.2 K-means Clustering. To determine whether susceptibility differences in the data could be categorized into distinct groups, we performed a cluster analysis using k-means clustering. Input to the clustering were the absolute and relative symptom mean values, i.e., two values per participant. To optimise its key parameter k—the number of clusters—we used the *NbClust* R package.[17] The package calculates 30 measures to determine the optimal k. As the similarity measure we used the *Euclidean distance* of these mean values. The majority of the indices (8) suggested using a cluster size of two. Therefore, we performed a *k-means* cluster analysis with two clusters, resulting in clusters with sizes of 28 and 43 with the within-cluster sum of squares of 33.09 and 22.42 and a between-cluster sum of square of 84.48. The distribution of participants to the susceptibility cluster is shown in Figure 16 (color). We found a mean absolute symptom score of $M_a = 2.04, SD = 0.50$ for the high susceptibility group, and a mean absolute symptom score of $M_a = 0.69, SD = 0.42$ for the low susceptibility group. For the relative symptom score, we found a mean of $M_r = 1.17, SD = 0.54$ for the high, and a mean of $M_r = 0.30, SD = 0.30$ for the low susceptibility group. In Table 11, we show a summary of the absolute and relative symptom scores. We found that 13 women and 16 men were categorized to the high-susceptible group, and 16 women and 27 men to the low-susceptible group. For the blue light user study, half of the participants (14) were categorized as low-susceptible, and half (14) as high-susceptible. The distribution for the two eye exercise studies was different. The first eye exercise study had 3 high-susceptible and 21 low-susceptible participants, and the second eye exercise study had 11 high-susceptible and 8 low-susceptible participants.

5.1.3 Three-way ANOVA. We determined the potential influence factors of **sex**, **susceptibility group**, and **user study** on DES mean values. We added the factor **user study** to the analysis, as potential differences in DES scores could be heavily influenced by the user study. On one hand, this could occur because the mean values for the participants of the three user studies were calculated differently. For the blue light study, the mean values were calculated based on one single item only, whereas they were calculated as a mean of 11 symptoms for the eye exercise user studies. On the other hand, although the experiments were similar in setup and duration, they differed in VR experience and rating scale. Summary statistics of the mean values are shown in Table 11.

We calculated a three-way ANOVA with three factors (levels are indicated in brackets) **sex (1)**, **susceptibility group (2)**, and **experiment (3)** and the dependent variables M_a and M_r. We

[17]RDocumentation NbClust: https://www.rdocumentation.org/packages/NbClust/versions/3.0/topics/NbClust, last retrieved: April 11, 2021.

Table 12. Summary Statistics of the Absolute and Relative Symptom Mean for Each Factor and Group

Susceptibility	Sex	Experiment	N	Variable	Mean	SD	Variable	Mean	SD
high	–	–	28	M_a	2.04	0.50	M_r	1.17	0.54
low	–	–	43	M_a	0.69	0.42	M_r	0.30	0.30
–	f	–	29	M_a	1.42	0.82	M_r	0.73	0.61
–	m	–	42	M_a	1.08	0.78	M_r	0.58	0.59
–	–	blue light study	28	M_a	1.31	0.88	M_r	0.90	0.66
–	–	1^{st} eye exercise study	24	M_a	0.72	0.50	M_r	0.26	0.27
–	–	2^{nd} eye exercise study	19	M_a	1.72	0.66	M_r	0.73	0.57

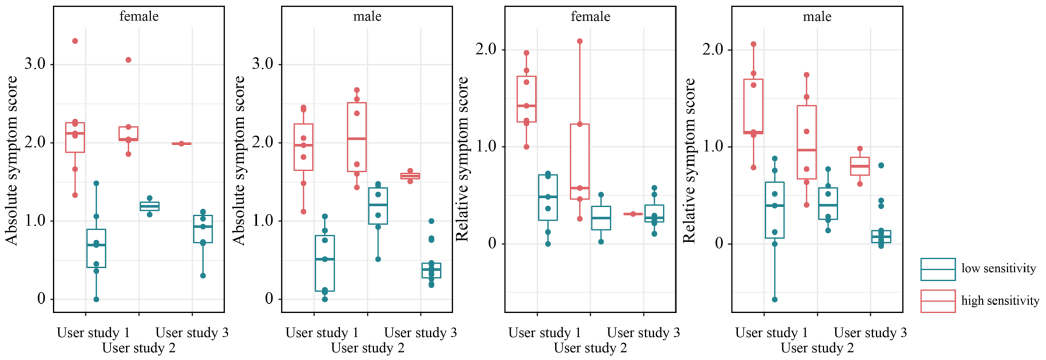


Fig. 17. Results of the three-way ANOVA for the absolute and relative symptom means of all participants. The three factors were **sex** (female, male), **susceptibility** (high, low), and **user study** (blue light user study (user study 1), 1^{st} eye exercise user study (user study 2), 2^{nd} eye exercise user study (user study 3).

checked the normality assumption by analyzing the residuals. We did not find a violation of this assumption by inspecting the QQ plots, where all points lay approximately on the reference line for both dependent variables M_a and M_r. This is supported by a non-significant Shapiro–Wilk normality test (M_a: $W = 0.99, p = .73$ and M_r: $W = 0.98, p = .30$). To test the assumption of homogenity of variances, we performed a Levene's test, which indicated that there is no violation of this assumption (M_a: $F(11, 59) = 1.09, p = .39$ and M_r: $F(11, 59) = 1.42, p = .19$). The 3-way ANOVA revealed a statistically significant small effect of *sex* ($F(1, 59) = 6.23, p = .02, eta_g^2 = .1$), a large effect of *susceptibility group* ($F(1, 59) = 120.25, p < .01, eta_g^2 = 0.67$), and a small effect of *experiment* ($F(2, 59) = 5.74, p < .01, eta_g^2 = 0.16$) on the absolute DES symptom mean M_a. For the relative DES symptom mean M_r, we found a statistically significant medium effect of *susceptibility group* ($F(1, 59) = 60.32, p < .01, eta_g^2 = 0.51$) and a small effect of *experiment* ($F(2, 59) = 4.55, p = .045, eta_g^2 = 0.13$). Boxplots of the results are shown in Figure 17.

For the factor **user study**, we conducted pairwise comparisons using a Wilcoxon test with Bonferroni correction. This test showed that there was a statistically significant difference between the blue light study and the first user study ($p = .05$) and between the first and the second user study ($p < .01$). The summary statistics for each factor are shown in Table 12.

5.1.4 *Discussion.*

On the Factors Sex and Susceptibility as Influences on DES. This analysis shows that the susceptibility to DES has a strong effect on DES symptoms. The effect of the factor susceptibility seems to be even stronger than the factor sex. Yet, women seem to be slightly more susceptible to experiencing more severe DES symptoms than men, as 45% of the female participants were attributed to

the high, and 55% to the low susceptibility group (N_{high} = 13, N_{low} = 16). For men, this was slightly different, with 37% of male participants belonging to the high, and 63% to the low susceptibility group (N_{high}=15, N_{low} = 26). However, it has to be considered that our overall sample was unbalanced with 41% female and 59% male participants. As the allocation to a cluster was performed on the basis of all participants, the results might be skewed toward male participants. Both the absolute and the relative DES symptoms were more severe for women than for men. However, this difference was only significant for the absolute DES symptom means. Our results indicate that the severity of symptoms varies between sexes, but that this effect is rather small. This is supported by the small effect size for the difference of the absolute values.

Furthermore, prior studies typically only reported pre vs. post-exposure values instead of differences in sensitivity of participants [63, 102]. For instance, Sharples et al. reported that 60–70% of their participants experienced an increase in symptoms over a time period of 30 minutes, but did not discuss differences between different groups of participants [84]. Our results suggest that especially the differences between susceptibility groups should be investigated in future works.

On the Factor User Study as Influence on DES. The statistically significant difference between the mean values of the three user studies (except the blue light study and the second user study on eye exercises) could be attributed to several factors. First, we used three different rating scales to assess symptoms, although, in all studies, we measured symptoms with 7-point scales. For the blue light user study, we used a rating scale ranging from *nothing at all* to *very severe*, asking participants to indicate how strong their perception of DES symptoms is right now. In the first user study on eye exercises, we asked participants to rate how strongly they perceived each symptom, using a scale ranging from *no problem* to *very severe problem*. Lastly, in the second user study on eye exercises, we used a scale ranging from *no discomfort* to *very severe discomfort*. While all three scales are unipolar scales grounded in 0 as not experiencing any symptoms and 6 experiencing very severe symptoms, the different labels and expressions might have contributed to differences in results. In addition to different rating scales, we also had two different VR experiences, which could have had an effect on DES symptoms. However, this is contradicted by the fact that there was a significant difference between the two user studies on eye exercises in which we used the same task. Furthermore, the different devices could have caused an effect on DES symptoms. Symptoms were lower with the Oculus Go headset (used in the first user study on eye exercises) compared to the Oculus Quest headset (used in the blue light study) and the HTC Vive (used in the second user study on eye exercises). Lastly, we conducted the three user studies in three different environments. While the blue light study and the first study on eye exercises were performed by participants in their homes, the second user study on eye exercises was conducted at a university. However, it is unlikely that the environment (being at home or not) had an effect on DES symptoms, as we observed differences in DES symptoms in the two studies that were conducted at home (blue light study and first study on eye exercises). Finally, as we aggregated the values per participant to one mean DES value for each participant, we might have lost a considerable amount of variance of the data. Yet, our work only lays the ground for a deeper analysis of different factors on DES across studies.

In general, the different distribution of high/low-susceptible participants across the three user studies is highly interesting. While in the blue light study, we found 50% high and 50% low-susceptible participants, the distribution was at 13%/87% for the first and 58%/42% for the second eye exercise study. In particular, the difference in the first eye exercise study stands out. Interestingly, this trend is reversed to what we would expect due to a habituation effect, as the sample in the blue light study was more experienced with VR-HMDs than the ones in the eye exercise studies. One reason could be that in the first eye exercise study, the sample contained proportionally

more men than women compared to the other studies. This, however, cannot entirely explain the big difference. We rather suspect that, by chance, we had more low-susceptible people in the sample than high-susceptible ones. Nonetheless, these outcomes underline the need for more future studies in this direction. Furthermore, the high number of low-susceptible participants in the first eye tracking study could explain why we could not observe significant differences between the eye exercises. With more high-susceptible participants in the second user study on eye exercises, the effect of these might just have been more observable as more participants experienced it.

5.2 Limitations and Key Challenges for Future Work

Key Challenge 1. Acknowledging DES as an important problem in VR HMDs and integrating its assessment in the evaluation of new VR-HMD experiences.

The results of the survey and user studies highlight that DES is a frequently occurring problem among VR-HMD users. Yet, compared to other forms of discomfort in VR-HMDs, such as simulator sickness, it is currently largely overlooked in the evaluation of new VR experiences [39]. With the increasing amount of research on VR-HMDs in the field of **human-computer interaction** (**HCI**) over the last years [37] and rising consumer adoption, DES will become increasingly important to VR-HMD researchers and users. As a first step to address the problem, researchers have to acknowledge its importance and, consequently, integrate its assessment in the evaluation catalog of new VR-HMD experiences and devices. Therefore, we suggest that DES should become a separate metric for evaluating future VR experiences and devices. While we understand that it may not be relevant for all VR experiences, its assessment should especially be considered in evaluating experiences that contain several straining factors, such as bright colors, high contrasts, or that require many unnatural eye movements.

Key Challenge 2. Investigating techniques to alleviate and prevent the problem of DES in VR HMDs.

In this article, we showed how potential alleviation techniques could be designed, implemented, and evaluated. While we focused on *alleviating* DES symptoms, the results of our second eye exercise study suggest that they could have an extended effect and thus potentially be used to *prevent* DES, too. A key challenge of future work will be to design both alleviation techniques but also prevention techniques.

Key Challenge 3. Developing alleviation and prevention techniques that are *integrated* into the VR experience.

It is not only important to investigate effective alleviation techniques but also to design them in a way that users want to apply them and do not feel disturbed by them in their VR experience. Based on the results of the eye exercise user studies, we propose some concepts for integrating alleviation techniques to reduce DES in VR-HMDs. *Closing the eyes* provides a fast and straightforward way to alleviate subjective DES in HMDs. One shortcoming is that it interrupts the users in

their VR experience. In contrast, eye exercises can be more fluently integrated into the VR experience, being short VR experiences themselves. However, their effectiveness is of shorter duration than closing the eyes. The integration of coping strategies into VR usage does not have to be a unique solution. Rather the combination of different solution approaches could lead to an overall more comfortable VR experience with a variety of techniques to address DES. In its simplest form, this can be a reminder to close the eyes for some time period, which would pose a VR pendant to the well-being functionalities integrated into smartphones that remind the user to take pauses. However, integrated eye exercises can be used even beyond this explicit way of asking the user to perform a task. Some eye exercises are designed around simple objects leading the user's gaze. This can easily be implicitly integrated into user interface elements (e.g., buttons that follow a similar motion) or even game mechanics (e.g., enemies flying toward the user). This could subconsciously trigger the necessary eye motions for the user and implicitly reduce DES in VR-HMDs. These could be fluently integrated into interface design with the user setting frequency and duration of occurrence. In a combined solution, integrated eye exercises that are constantly present during a VR experience could be used to mitigate symptoms at short notice, while an additional reminder suggests users close their eyes every 30 minutes of usage to reduce extended effects. In summary, alleviation techniques should be used complementary, leading to an overall more comfortable and less straining VR experience.

Key Challenge 4. Considering a person's susceptibility to DES in the design of new alleviation and prevention techniques.

DES is a problem that increases over time and does not occur in a similar manner for all users. Some users experience strong symptoms shortly after initiating a new VR experience, while others experience symptoms only after a prolonged time of use. Designers should, therefore, think of alternative solutions for straining tasks according to the individual needs of users. Whereas for a person that is not particularly sensitive to DES, an (eye-based) interaction strategy might not have negative effects on their user experience or cause more symptoms, the same technique, when applied by a high-sensitivity person might worsen their experience significantly by leading to an increase in symptoms. This is already compensated for in simulator sickness, where users get the option to either use smooth locomotion (for people that do not experience simulator sickness strongly) or teleportation (for people that experience more simulator sickness). In a similar fashion, alternatives for different types of users should be implemented for HMDs.

Key Challenge 5. Considering that unnatural gaze behavior might worsen DES in the design of explicit gaze-based interaction techniques.

Our findings are particularly important for the design of new gaze-based interaction techniques for VR-HMD experiences. In particular, with eye tracking being increasingly integrated into VR-HMDs, it should be investigated carefully whether potential eye-based interaction techniques pose significant additional strain to users' eyes. It was shown that prolonged fixation duration and many large saccades could increase DES [38]. Therefore, it should be avoided to include these eye movements in eye-based interaction strategies. On the other hand, eye-based interaction techniques could specifically be designed to alleviate DES symptoms. Our eye exercises are one

example of an interaction technique that aims at reducing DES actively. Other eye-based interaction techniques could be invented that passively cause less DES. In the design of eye-based interaction techniques, we should consider that inducing unnatural gaze behavior can both increase and reduce DES. For example, while it may be beneficial to users to rest their gaze on an interactive element for a while when exposed to a reading task (as applied in our study), in other cases, where the interaction relies on prolonged fixation duration, dynamic eye movements can be used to alleviate strain. Therefore, it is important to study the effects of different eye movements in different contexts.

Key Challenge 6. Finding a good trade-off between detailed symptomatology and high temporal resolution when measuring DES in VR-HMDs.

In our studies, we ran a two-track model to measure DES by collecting complete questionnaires and single items of DES. Using this strategy, we could obtain detailed answers about users' symptomatology using a complete questionnaire directly after the exposure, but also a high temporal resolution using a single item of DES during the exposure. While in the eye exercise studies, we measured 11 symptoms of DES during the conditions, in the blue light study, we only measured one summary DES item during the exposure. While we hoped to get more detailed insights using the complete set of symptoms during the exposure in the eye exercise studies, we found that it is more beneficial if participants have to answer only one single item question repeatedly. Therefore, we recommend the following strategy to obtain DES symptoms in VR experience: measuring detailed symptomatology of DES after the exposure and focus on high temporal resolution when measuring DES during the exposure by applying one single item repeatedly. The drawback of this technique is that users have to be interrupted in their experience to measure symptoms. In the blue light study, we designed the VR experiences in a way that the measurement can be integrated between levels. However, this might not always be possible, e.g., if the VR experience cannot be subdivided into several levels. Therefore, passive or objective measures would be helpful to measure DES. Although literature showed that DES could be measured objectively by eye tracking metrics, [97], we could not observe effects for objective measures in our second user study on eye exercises. One reason for this might be that we could not constantly receive high-quality data from the eye tracker, or that, despite works that have used eye tracking to detect DES objectively [30, 73, 89], more specific optometric measures are needed. The measurement of objective DES indicators also remains a challenge, as gaze metrics, such as pupil size or blink rate, are closely coupled to cognitive and affective responses [70, 92] and can only be applied in settings that are specifically tailored to measure DES [97], but do not necessarily reflect on natural usage behavior.

Key Challenge 7. Developing a suitable rating scale to assess the range of severity of symptoms that is relevant to users and researchers.

The rating scale that reflects users' current state of symptoms best remains an open question. In our three user studies, we used 7-point symptom severity scales anchored in the value of 0 that indicates that users are not experiencing any symptoms. The maximum of the scales was labeled as *experiencing very severe symptoms or discomfort*. However, it is questionable whether this type of scale covers the relevant symptom range for VR researchers in sufficient detail. In designing VR experiences, we have to ask ourselves what the threshold is that we consider an acceptable value

of DES. In our investigations on DES, we experienced that users are in general affected, but rarely in a very severe way. Therefore, it might be beneficial to change the maximum pole of the rating scale to a lower value, such as "moderate", which is already considered an intensity that we do not want users to experience after a 25-minute exposure to a VR-HMD. This would then allow a more finely granular scale, with more levels in the lower intensity ranges, which would provide a more detailed picture of the perceived DES symptoms.

Key Challenge 8. Investigating further influence factors that drive DES in VR-HMDs, including a more detailed investigation into the factor susceptibility.

In our final analysis, we found two influence factors (*sex* and *susceptibility*) that particularly drive DES in VR-HMDs. With our analysis, we showed that these factors are indeed distinct for DES, while previous works have suggested that they are related to each other for simulator sickness [21, 33]. Yet, we have to consider that our analysis was conducted on a set of three different user studies and that we had a sample with more men (59%) than women (41%). Furthermore, the distinction of participants into two groups was based on the k-means clustering algorithm that was performed on the same set of participants. To achieve a more generic statement on susceptibility, we should investigate whether this distinction can be made for a larger set of participants and whether participants that are not part of the "definition set" can also be classified as high or low-susceptible. Furthermore, future work has to investigate whether the definition of susceptibility can be determined based on an absolute value (e.g., the CVS-Q suggests a value of six with which users are classified as suffering from the CVS [28]). For now, we did not determine susceptibility based on one specific value but divided the set of participants into two sets based on their averaged absolute and relative DES scores. However, this separation is not final, and it remains an open question whether participants can be grouped into more than two susceptibility groups. Researchers should be aware of the two factors, sex and susceptibility, in their sample, as this could heavily influence the effect of potential alleviation techniques (e.g., when having a sample that is in general low-susceptible, the effects of potential alleviation techniques would be difficult to reveal).

As we discuss in Section 5.1.4, we expect that there are more factors that can have an influence on DES in VR-HMDs that are not yet revealed completely. It seems obvious that these include the task or type of experience. While we know that reading, for instance, worsens DES [25, 86], further investigations are needed to determine VR-specific influence factors. The majority of articles that investigate influence factors on DES during natural behavior were conducted with conventional computer displays. However, in contrast to VR-HMDs, these displays do not cover the entire field of view. For example, it could be possible that factors such as movement in the periphery during first-person shooter games in VR might affect DES by causing more saccades to the outer field of view. Such VR-specific factors are currently unknown. Finally, ranking different influence factors is a fundamental challenge for future work to focus research efforts on the most important ones.

Key Challenge 9. Investigating DES effects that occur after repeated exposure, long-term usage, or longer time periods of use.

VR-HMDs have only started to enter the consumer market, and studies with VR-HMDs are typically conducted in laboratory settings, with users being exposed to the technology only a few times and for a limited time period. With our first study on eye exercises, we exposed the participants to the devices repeatedly over a period of 10 days. While significantly longer than previous studies [58, 72, 84], this still only provides a glimpse into natural use of the device in private life (repeated exposures for weeks, months, or even years). Furthermore, while this study was conducted with users at home, the application that they had to play was a given study application. Further studies with users at home have to be conducted to gain insights into the symptomatology of users during natural behavior. We gained a generic overview of these symptoms with our survey, but more detailed insights should be further investigated. In all three user studies, we found that the severity of DES symptoms increased significantly over the 25 minutes of exposure. We do currently not know how a repeated experience of these symptoms over a longer time period of using the devices might impact users. It is important to investigate whether repeated exposure leads to a habituation effect and users would stop feeling symptoms after a certain time period of using the device. On the other hand, it could well be that symptoms get stronger over time when being repeatedly exposed to the technology.

Furthermore, long-term effects are currently unknown. VR-HMDs have only started to being used in an every-day fashion. Similar to long-term effects of smartphones, where long-term usage is associated with changes in refractive error [32], similar effects could occur with VR-HMDs, especially as they are put even closer to the eyes.

Lastly, DES symptoms have typically been tested for a limited amount of time (around 30 minutes). Using the devices for longer time periods might change the experience of symptoms from temporal to consistent. Overall, the wider distribution of the devices to consumer users gives us the opportunity and responsibility to conduct such long-term investigations in the near future.

Key Challenge 10. Investigating the relationship of DES to other usability metrics in VR, such as presence or enjoyment.

The experience of DES in VR-HMDs might have influences on other usability metrics of VR-HMDs, such as presence or enjoyment. For example, the literature suggests that simulator sickness and presence are negatively correlated [48]. Similar investigations have not been conducted for DES. In our studies, we could not observe statistically significant differences in presence or enjoyment values between the conditions. However, we did not specifically aim at investigating such an effect, and, therefore, our study designs were not tailored toward investigating this relation. Such a relationship would have important impacts on study designs. For instance, more severe symptoms of DES might distract users from feeling present in the virtual environment. On the other hand, an increased feeling of presence might lead to a decrease in the perception of symptoms, and we have to investigate whether this would be desirable.

6 CONCLUSION

In this work, we presented the first comprehensive investigation into DES in VR-HMDs. We presented the results of an online survey with 68 experienced VR-HMD users, revealing details about symptomatology. The lack of integrated solutions, as reported by the online survey respondents, motivated us to design and evaluate two alleviation approaches in three user studies (N = 71). Blue light filtering, being the first approach, did not show a positive effect on DES symptoms. However, eye exercises, when applied for a duration of 2:30 minutes, effectively reduced DES symptoms.

Furthermore, we conducted an analysis of all three user studies, which revealed *sex* and *susceptibility* as two important factors that drive DES in VR-HMDs. Lastly, we summarized the findings of these three parts (*understanding*, *addressing*, and *analyzing*) into 10 key challenges for future research on DES in VR-HMDs. Most importantly, DES has to be acknowledged as a severe problem in VR-HMD use by the research community. We, therefore, argue that DES should become an essential part of evaluating VR-HMD devices and experiences.

APPENDICES

A UNDERSTANDING DES IN VR-HMD USERS: ONLINE SURVEY QUESTIONNAIRE

A.1 Demographics

Question 1: How old are you?
Question 2: To which gender identity do you most identify?
- Woman
- Man
- Non-binary
- Prefer not to disclose
- Prefer to self-describe

Question 3: What is the highest degree or level of school you have completed?
- No schooling completed
- High school degree or equivalent
- Bachelor's degree (e.g. BA, BS)
- Master's degree (e.g. MA, MS, MEd)
- Doctorate (e.g. PhD, EdD)

Question 4: Are you currently...?
- Employed for wages
- Self-employed
- Unemployed
- Homemaker
- Student
- Retired
- Unable to work

A.2 Devices and Usage

Question 1: Do you own or use a VR headset regularly (once a week or more often)?
- I've used a VR headset regularly in the past, but not at the moment
- I've tried VR, but never used a headset regularly
- No, I've never used a VR headset

Question 2: Which of the following VR headsets do you use regularly (i.e., once a week or more often)? If you use more than one headset, please select the one you use most often.
- Google Carboard
- Samsung Gear VR
- Oculus Go
- Sony PlayStationVR
- Google Daydream View
- HTC Vive, HTC Vive Pro, HTC Vive Focus, or HTC Vive Pro Eye
- Oculus Rift, Oculus Rift S, or Oculus Rift (DK1/DK2)
- Oculus Quest

— Valve Index
— Samsung Odyssey
— Nintendo Labo VR

Question 3: How often do you use the VR headset?
— several times a day
— once a day
— once or twice a week
— once a week

Question 4: For how long have you actively been using the VR headset regularly (i.e., once a week or more often)?
— less than 1 month
— 1–6 months
— 6–12 months
— 2 years
— more than 2 years

Question 5: For how long do you usually use the VR headset in one session?
— less than 30 minutes
— 30–60 minutes
— 1–2 hours
— 2–3 hours
— more than 3 hours

Question 6: In the following please indicate how often you usually interrupt a VR session in one hour (i.e., taking off the headset)...

Question 7: Visual discomfort hereby refers to symptoms affecting the eyes due to the exposure to digital screens. Symptoms include blurred vision, pain around the eyes, headache, or dry eyes. Simulator Sickness is a form of motion sickness occurring in VR. Symptoms include nausea, vomiting, sweating, headaches, uneasiness, drowsiness, and disorientation.

	never	once	twice	3 times	4 times	5 times	more often
... in general?	o	o	o	o	o	o	o
... due to symptoms of visual discomfort?	o	o	o	o	o	o	o
... due to symptoms of simulator sickness?	o	o	o	o	o	o	o
... due to symptoms of general discomfort?	o	o	o	o	o	o	o
...other (e.g., to use the smartphone, to go to the bathroom)	o	o	o	o	o	o	o

Question 8: Please enter a number for how often you interrupt a VR session in general.

Question 9: Please enter a number for how often you interrupt a VR session due to symptoms of visual discomfort.

Question 10: Please enter a number for how often you interrupt a VR session due to symptoms of simulator sickness.

Question 11: Please enter a number for how often you interrupt a VR session due to symptoms of general discomfort.

Question 12: When did you last use the VR headset?
— just now (a few minutes ago)
— recently (a few hours ago)
— yesterday
— days ago

— weeks ago
— months ago

Question 13: Do you have a VR headset that allows for adjusting the interpupillary distance (e.g., hardware slider or software based)?

Question 14: Interpupillary distance refers to the distance between the centers of the pupils of the eyes.
— yes
— no
— I don't know

Question 15: How often do you adjust the interpupillary distance?
— Once in the beginning.
— When I'm noticing some issues, such as blurred vision or double vision.
— Most of the times when I use the device.
— Everytime I use the device.
— Never.

Question 16: Have you ever noticed any problems due to incorrect settings of the interpupillary distance?
— Blurred vision
— Double vision
— General visual discomfort
— I don't know
— No, I haven't noticed any problems

A.3 Eye Specific Demographics

Question 1: Do you have any vision problems, such as ...
— Nearsightedness (things far away are difficult to see clearly)
— Farsightedness (close things are difficult to see clearly)
— Astigmatism (cornea is irregularly shaped)
— Strabismus (Crossed Eyes)
— No, I don't have any vision problems.

Question 2: Do you regularly wear prescription glasses or contact lenses to correct your vision problems?
— Yes, I use both, prescription glasses and contact lenses.
— Yes, I use prescription glasses.
— Yes, I use contact lenses.
— No, I don't use either.

Question 3: When do you wear prescription glasses?
— All the time
— When using a computer
— To read
— When driving
— When doing sports
— Other

Question 4: When do you wear contact lenses?
— All the time
— When using a computer
— To read
— When driving

— When doing sports
— Other

Question 5: Do you use your vision aids (prescription glasses or contact lenses) when using VR? (If you use both please indicate the dominant form of vision aids you use in VR.)
— I wear prescription glasses when using VR.
— I wear contact lenses when using VR.
— No, I don't use vision aids when using VR.

Question 6: Do you have any type of visual impairment, such as ...
— Cataract
— Glaucoma
— No, I don't.
— Other

A.4 General DES Experience

Question 1: In general, please indicate whether you usually experience any form of visual discomfort ...

	never (the symptom does not occur at all)	**occasionally** (sporadically or once per hour of usage)	**often or always** (2 or 3 times per hour of usage or almost every time)
during a VR session.	o	o	o
after a VR session.	o	o	o

Question 2: Please specify the intensity of visual discomfort ...

	none	moderate	intense
during a VR session.	o	o	o
after a VR session.	o	o	o

Question 3: After what period of time does visual discomfort on average occur **during** a VR session?

Question 4: After what period of time does visual discomfort on average occur **after** a VR session?

Question 5: How quickly do the symptoms subside?
— Within seconds after usage.
— Within 5 minutes after usage.
— Within 15 minutes after usage.
— Within 30 minutes after usage.
— Within an hour after usage.
— It takes more than an hour after usage.
— Other

Question 6: How does the occurrence of visual discomfort compare to looking at a desktop monitor for the same amount of time?
— I experience visual discomfort **more frequently** when using a **desktop monitor**.
— I experience visual discomfort **more frequently** when using a **VR headset**.
— I experience visual discomfort **about equally as often** when using a VR headset and a desktop monitor.

Question 7: Do you remember one specific scenario where you experienced visual discomfort? If yes please indicate how and why it occurred and what you did to mitigate the symptoms.

Question 8: Do you have any further comments on the occurrence of visual discomfort using VR headsets?

A.5 Computer Vision Syndrome-Questionnaire

Question 1: Please indicate whether you experience any of the following symptoms during the time you use a VR headset (with regards to your eyes). If you select "never" on the FREQUENCY scale, please select "N/A" on the INTENSITY scale.

	a. FREQUENCY			b. INTENSITY		
	(the symptom does not occur at all) NEVER	(sporadic episodes or once per hour of usage) OCCASIONALLY	(2 or 3 times per hour of usage or almost every time) OFTEN OR ALWAYS	MODERATE	INTENSE	N/A
Burning	o	o	o	o	o	o
Itching	o	o	o	o	o	o
Feeling of a foreign body	o	o	o	o	o	o
Tearing	o	o	o	o	o	o
Excessive blinking	o	o	o	o	o	o
Eye redness	o	o	o	o	o	o
Eye pain	o	o	o	o	o	o
Heavy eyelids	o	o	o	o	o	o
Dryness	o	o	o	o	o	o
Blurred vision	o	o	o	o	o	o
Double vision	o	o	o	o	o	o
Difficulty focusing for near vision	o	o	o	o	o	o
Increased sensitivity to light	o	o	o	o	o	o
Coloured halos around objects	o	o	o	o	o	o
Feeling that sight is worsening	o	o	o	o	o	o
Headache	o	o	o	o	o	o

A.6 Coping Strategies

Question 1: Are you aware that coping strategies or eye exercises to reduce visual discomfort exist?

— Yes

— No

Question 2: Do you know and use any of the following coping strategies or eye exercises?

	I know and use that strategy.	I do that, but I wasn't aware that it is a strategy.	I know that strategy, but I don't use it.	No, I don't know that strategy.
Looking at something 20 feet away for20 seconds every 20 minutes (20-20-20 rule)	o	o	o	o
Squinting the eyes for a while	o	o	o	o
Blinking quickly for a while	o	o	o	o
Closing the eyes for a while	o	o	o	o
Covering the closed eyes with the hands to relax the eyes (palming)	o	o	o	o
Rolling the eyes into one direction for a while	o	o	o	o

Question 3: Do you use a strategy that isn't listed here?
- No, I don't.
- Yes, the following:

Question 4: How often and for how long do you use a strategy/do eye exercises during a VR session? Please choose:
Frequency:
- never
- less than once an hour
- once an hour
- 2 an hour
- 3 times an hour
- more than 3 times an hour

Duration:
- never
- for 10 seconds.
- for 20 seconds.
- for 30 seconds.
- for 1 minute.
- for 2 minutes.
- for 5 minutes.
- for more than 5 minutes.

Question 5: If you are aware that strategies to prevent visual discomfort exist, but don't use them - please indicate why.
- I forget about them.
- It's too much effort/cumbersome.
- It would be too interruptive.
- Doesn't deem necessary (symptoms go away by themselves after a while).
- Other:

Question 6: Do you use an eye exercises or eye training mobile app?
- No, I don't.
- Yes, the following:

Question 7: Do you have any further comments regarding strategies to prevent or mitigate visual discomfort?

A.7 Usage of Coping Strategies

Question 1: Now that you are aware that visual discomfort may occur after longer exposure to VR and that certain strategies exist to mitigate the symptoms.

Question 2: Please rate based on agreement/disagreement:

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
I would be willing to remove the headset to perform an eye exercise.	o	o	o	o	
I would be willing to perform an eye exercise in VR	o	o	o	o	
I would be willing to perform eye exercises that are implicitly integrated in the VR experience, e.g., in loading screens.	o	o	o	o	

Question 3: I would be willing to remove the headset to perform an eye exercise ...
Frequency

— never
— less than once every 3 hours
— once every 3 hours
— once every 2 hours
— one an hour
— 2 times an hour
— 3 times an hour
— more than 3 times an hour

Duration
— never
— for 10 seconds.
— for 20 seconds.
— for 30 seconds.
— for 1 minute.
— for 3 minutes.
— for 5 minutes.
— for more than 5 minutes

Question 4: I would be willing to perform an eye exercise in VR ...

Frequency
— never
— less than once every 3 hours
— once every 3 hours
— once every 2 hours
— one an hour
— 2 times an hour
— 3 times an hour
— more than 3 times an hour

Duration:
— never
— for 10 seconds.
— for 20 seconds.
— for 30 seconds.
— for 1 minute.
— for 3 minutes.
— for 5 minutes.
— for more than 5 minutes

Question 5: Please rate based on agreement/disagreement: I am concerned that the use of a VR headset will negatively impact my eyes (e.g., eye health, wellbeing, and sight).
— Strongly Disagree
— Disagree
— Neither Agree nor Disagree
— Agree
— Strongly Agree

Question 6: Do you have any further comments regarding the usage of strategies to prevent or mitigate visual discomfort while using a VR headset?

B APPLYING BLUE LIGHT FILTERING TO ADDRESS DES IN VR HMDS (USER STUDY 1)

B.1 Pre-study Experiments to Determine the Properties of the Peripheral Filter

As the peripheral filter has to be implemented in a way that users would accept and use it, we evaluated its intensity values in a pre-study consisting of three experiments with seven participants (two female, five male, $M = 22, SD = 3$). The radius of the filter was set to 25° from the viewing centre. We approached this problem from two sides. First, we determined a detection threshold of the peripheral filter, i.e., at which percentage of blocked light participants would recognize that a filter is active. In the second experiment, we aimed at determining an acceptance threshold of the peripheral filter, i.e., at which percentage the participants would accept to use a strong filter. Lastly, we asked participants for their personal preferences about the filters. The first experiment was implemented as a two-alternative-forced-choice design, which has previously been used to determine perceptual effects [77, 78]. Using this approach one can determine the probability of how well participants are able to recognize an active filter. In each iteration, participants were presented with two versions of the filter, where only in one version the filter was active. If they cannot distinguish between the filters, a probability of 50% will be reached. We tested three versions of the filter, blocking 20%, 30%, and 40% of blue light, which results in eight combinations for the experiment. Each combination was repeated eight times, resulting in a total iteration count of 48. For each iteration, participants played a VR game two times, each for 15 seconds (*LightSaber* Section 4.1.4). After each iteration, they had to decide in which of them a filter was active.

In the second experiment, we aimed at determining to what extent strong filters were accepted by users. To determine the upper limit of acceptance, we implemented again three versions of the filter, blocking 60%, 80%, and 100% of the blue light. As the perception of a certain filter is probably influenced by the order of presentation (e.g., the perception of a lighter filter might be influenced by the previous perception of a strong filter), we decided to let participants rate the three filter versions both in ascending and in descending order. To avoid further influences, half of the participants were presented with ascending order first, and half of the participants were presented with the descending order first. Participants played the LightSaber VR game for 60 seconds for each of the filters. After each iteration, they answered a set of 11 7-point Likert scale questions, regarding the perception of the filter.

B.1.1 Measures. Questions that were assessed during the pre-study experiment to determine the acceptance threshold of the peripheral filter. The questions were answered on a 7-point Likert scale, ranging from *strongly disagree* to *strongly agree*.

- It feels natural to play with a filter like this in Virtual Reality.
- I think the presentation of colors is not disturbed when using this of filter.
- I would personally use a filter like this to take care on my eyes when using Virtual Reality.
- Using this filter negatively influences my playing experience.
- I like the aesthetics of this filter.
- Using this filter is pleasant.
- I think the presentation of colors is disturbed when using this of filter.
- I feel present in the virtual environment, as if I were really there.
- Using this filter doesn't influence my playing experience.
- Using this filter is pleasant for my eyes.
- I very well recognized this filter.

Lastly, we asked participants to adjust a peripheral blue light filter according to their preferences while playing the LightSaber game for six minutes. At, this participants were asked to alter the

filter intensity in discrete steps until they reached a level that felt comfortable to them and that they would still accept. Participants did this two times, one in ascending order and once in descending order (each for three minutes).

B.1.2 Results. In the first experiment, we measured an average error rate of 34% ($SD = 0.15$) for the 20% filter, an error rate of 29% ($SD = 0.14$) for the 30% filter, and an error rate of 23% ($SD = 0.21$) for the 40% filter. For the second experiment, the majority of participants gave positive feedback about the 60% and the 80% filter. They even found that the appearance of colors is not disturbed when using these particular filters, although they recognized the filter well. For the 100% filter, participants stated that they clearly recognized it, but still they stated that their playing experience was not negatively influenced. Lastly, for the personal preference, we found a median of 60% for the incremental and the decremental filter adjustment. These results motivated us to use a strong filter in the study, as the user experience did not seem to be too negatively influenced by a strong filter. Furthermore, we expected to measure larger effects when using a stronger filter.

B.2 Baseline Calculation of Intermediate Questions

We computed a baseline value for each participant by calculating the mean of the first three values ($mean(Q_0, Q_1, Q_2)$). This mean was subtracted from all values Q_0–Q_{10}. By doing so, we eliminated potential effects that could occur when participants experienced symptoms already before starting the study. Furthermore, we eliminated potential familiarization effects of filling out the questionnaire for the first time. To determine the baseline, we proceeded as follows. For each participant and each condition, we first determined the position of the first increase in symptoms (Q_0 when participants did not experience any changes in symptoms). We then calculated the median of the positions of the first increase for all participants and all conditions. The median of this first increase in DES score was Q_2. Therefore, we chose the mean of Q_0, Q_1, and Q_2 as the baseline value.

B.3 Pre/Post Condition Discomfort Questionnaire for the Blue Light Study

Table 13. Symptoms of DES and Simulator Sickness That Were Measured in the Pre and Post-condition Questionnaires in the Blue Light User Study

Type	Question	Scale
Digital eye strain	Blurred vision	7-point Likert scale Nothing at all (I don't experience this at all) - Very severe (I don't want to use the device under these conditions)
Digital eye strain	Burning eyes	
Digital eye strain	Difficulty concentrating	
Digital eye strain	Difficulty focusing	
Digital eye strain	Dry eyes	
Digital eye strain	Eye redness	
Digital eye strain	Eye strain	
Digital eye strain	Excessive blinking	
Digital eye strain	Feeling of a foreign body	
Digital eye strain	Feeling that sight is worsening	
Digital eye strain	Heavy eyelids	
Digital eye strain	Increased sensitivity to light	
Digital eye strain	Irritated eyes	
Digital eye strain	Neck pain	
Digital eye strain	Seeing colored halos around objects	
Digital eye strain	Sensation of hot eyes	
Digital eye strain	Shoulder pain	
Digital eye strain	Soreness of eyes	
Digital eye strain	Tearing eyes	
Digital eye strain	Tired eyes	
Digital eye strain	Watering of eyes	
Simulator sickness	Blurred vision	7-point Likert scale Nothing at all (I don't experience this at all) - Very severe (I don't want to use the device under these conditions)
Simulator sickness	Burping	
Simulator sickness	Difficulty concentrating	
Simulator sickness	Difficulty focusing	
Simulator sickness	Eye strain	
Simulator sickness	Dizziness with eyes closed	
Simulator sickness	Dizziness with eyes open	
Simulator sickness	Fatigue	
Simulator sickness	"Fullness of head"	
Simulator sickness	General discomfort	
Simulator sickness	Headache	
Simulator sickness	Nausea	
Simulator sickness	Salivation increasing	
Simulator sickness	Stomach awareness (is usually used to indicate a feeling of discomfort which is just short of nausea.)	
Simulator sickness	Sweating	
Simulator sickness	Vertigo (is experienced as loss of orientation with respect to vertical upright.)	

Table 14. Ergonomic Symptoms That Were Measured in the Post-condition Questionnaires in the Blue Light User Study

Type	Question	Scale
Ergonomic Symptoms	I felt tense or on edge because I was wearing the device.	7-point Likert scale Nothing at all (I don't experience this at all) - Very severe (I don't want to use the device under these conditions)
Ergonomic Symptoms	I felt that I did not have the device properly attached.	
Ergonomic Symptoms	I felt bulky wearing the device.	
Ergonomic Symptoms	I felt strange wearing the device	
Ergonomic Symptoms	I did not feel safe wearing the device.	
Ergonomic Symptoms	The device was painful to wear.	
Ergonomic Symptoms	The attachment of the device was too loose.	
Ergonomic Symptoms	The device generated additional heat leading to excess sweating.	
Ergonomic Symptoms	Wearing the device made me feel physically different.	
Ergonomic Symptoms	The device inhibited or restricted my movement	
Ergonomic Symptoms	I could feel the device on my body.Heavy eyelids	
Ergonomic Symptoms	The attachment of the device was too tight.	
Ergonomic Symptoms	I was worried about how I look when I wear this device.	
Ergonomic Symptoms	I felt the device was too heavy.	
Ergonomic Symptoms	The device was causing me some harm.	
Ergonomic Symptoms	I was not able to move as usual	
Ergonomic Symptoms	The device affected the way I move.	
Ergonomic Symptoms	I could feel the device moving.	
Borg10 Scale	Forehead	7-point Likert scale Nothing at all (I don't experience this at all) - Very severe (I don't want to use the device under these conditions)
Borg10 Scale	Temples	
Borg10 Scale	Cheeks	
Borg10 Scale	Nose	
Borg10 Scale	Ears	
Borg10 Scale	Eyes	
Borg10 Scale	Back of the head	
Borg10 Scale	Neck	
Borg10 Scale	Shoulders	

B.4 Usability Questionnaire for the Main Study

Table 15. Usability Questions That Were Asked in the Post-condition Questionnaire in the Blue Light Study

Type	Question	Scale
Visual appeal (PXI [1])	I enjoyed the way the game was styled.	
Visual appeal (PXI [1])	I liked the look and feel of the game.	
Visual appeal (PXI [1])	I appreciated the aesthetics of the game.	
Enjoyment (PXI [1])	I enjoyed playing the game.	7-point Likert scale
Enjoyment (PXI [1])	I liked playing the game.	strongly agree - strongly disagree
Enjoyment (PXI [1])	Playing the game was fun.	
Enjoyment (PXI [1])	The game was entertaining.	
Enjoyment (PXI [1])	I had a good time playing this game.	
IPQ [83]	Somehow I felt that the virtual world surrounded me.	
IPQ [83]	I felt like I was just perceiving pictures.	
IPQ [83]	I had a sense of acting in the virtual space	
IPQ [83]	I felt present in the virtual space.	7-point scale
IPQ [83]	I was not aware of my real environment.	fully disagree - fully agree
IPQ [83]	I still paid attention to the real environment.	
IPQ [83]	I was completely captivated by the virtual world.	
IPQ [83]	The virtual world seemed more realistic than the real world.	
IPQ [83]	In the computer generated world I had a sense of "being there".	7-point scale not at all - very much
IPQ [83]	I did not feel present in the virtual space.	7-point scale did not feel - felt present
IPQ [83]	How aware were you of the real world surrounding while navigating in the virtual world? (i.e., sounds, room temperature, other people, etc.)?	7-point scale extremely aware - not aware at all
IPQ [83]	How real did the virtual world seem to you?	7-point scale about as real as an imagined world - indistinguishable from the real world
IPQ [83]	How much did your experience in the virtual environment seem consistent with your real world experience?	7-point scale not consistent - very consistent
IPQ [83]	How real did the virtual world seem to you?	7-point scale not real at all - completely real
Obtrusiveness	The visuals of the game felt natural.	
Obtrusiveness	The colors of the game looked distorted.	7-point Likert scale
Obtrusiveness	The visuals of the game negatively influenced my experience.	strongly agree - strongly disagree
Obtrusiveness	Playing the game felt pleasant for my eyes.	
Obtrusiveness	I noticed that some visual overlay was present.	

B.5 Cube Spawn Positions of *LightSaber* Game

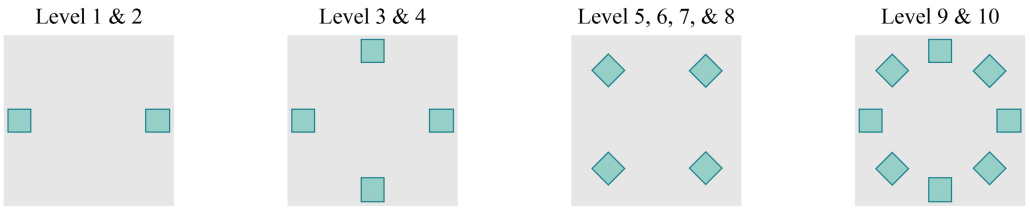


Fig. 18. Spawn positions of the cubes in the VR game *LightSaber* that was implemented for the blue light study.

B.6 List of Songs

The songs used in the blue light study are by "TheFatRat" https://www.youtube.com/user/ThisIsTheFatRat.

The titles used are:

(1) Xenogesis: https://www.youtube.com/watch?v=3_-a9nVZYjk
(2) Fly Away: https://www.youtube.com/watch?v=cMg8KaMdDYo
(3) No No No: https://www.youtube.com/watch?v=d0uFvhCHWCo
(4) Close to the Sun: https://www.youtube.com/watch?v=oJuGlqO85YI
(5) Epic: https://www.youtube.com/watch?v=AgPbZHXQNAU
(6) Chosen: https://www.youtube.com/watch?v=9YHTVML4PTE
(7) Jackpot: https://www.youtube.com/watch?v=kL8CyVqzmkc
(8) Unity: https://www.youtube.com/watch?v=n8X9_MgEdCg
(9) The Calling: https://www.youtube.com/watch?v=KR-eV7fHNbM
(10) Timelapse: https://www.youtube.com/watch?v=3fxq7kqyWO8
(11) Monody: https://www.youtube.com/watch?v=B7xai5u_tnk
(12) Solitude: https://www.youtube.com/watch?v=cIYdfWFsMXw

B.7 Statistical Results for the Main Study

B.7.1 Shapiro-Wilk Normality Tests.

Table 16. Shapiro–Wilk Normality Test Results for the Two-factor Analysis of the Within-condition Questions

Condition	Time	Variable	Test statistic	p-value
global filter	Q_0	relative DES score	0.7500531	1.580995e-05
global filter	Q_1	relative DES score	0.8160829	2.078621e-04
global filter	Q_2	relative DES score	0.8341135	4.541997e-04
global filter	Q_3	relative DES score	0.8420467	6.486697e-04
global filter	Q_4	relative DES score	0.8314844	4.042991e-04
global filter	Q_5	relative DES score	0.8317361	4.088147e-04
global filter	Q_6	relative DES score	0.8791273	3.841417e-03
global filter	Q_7	relative DES score	0.9155262	2.692638e-02
global filter	Q_8	relative DES score	0.9082339	1.793065e-02
global filter	Q_9	relative DES score	0.9060596	1.590817e-02
global filter	Q_{10}	relative DES score	0.9130630	2.345030e-02
peripheral filter	Q_0	relative DES score	0.8185916	2.312213e-04
peripheral filter	Q_1	relative DES score	0.8332034	4.362226e-04
peripheral filter	Q_2	relative DES score	0.8623669	1.678260e-03
peripheral filter	Q_3	relative DES score	0.8650929	1.914843e-03
peripheral filter	Q_4	relative DES score	0.9124936	2.271582e-02
peripheral filter	Q_5	relative DES score	0.9191239	3.300080e-02
peripheral filter	Q_6	relative DES score	0.9356303	8.558918e-02
peripheral filter	Q_7	relative DES score	0.9389315	1.038281e-01
peripheral filter	Q_8	relative DES score	0.9529697	2.350686e-01
peripheral filter	Q_9	relative DES score	0.9515737	2.169666e-01
peripheral filter	Q_{10}	relative DES score	0.9678883	5.252526e-01
control condition	Q_0	relative DES score	0.6943593	2.373872e-06
control condition	Q_1	relative DES score	0.7314953	8.202340e-06
control condition	Q_2	relative DES score	0.6451718	5.227773e-07
control condition	Q_3	relative DES score	0.8009601	1.109926e-04
control condition	Q_4	relative DES score	0.8789659	3.810158e-03
control condition	Q_5	relative DES score	0.8822376	4.500360e-03
control condition	Q_6	relative DES score	0.8828571	4.645355e-03
control condition	Q_7	relative DES score	0.9268345	5.132640e-02
control condition	Q_8	relative DES score	0.9344043	7.967322e-02
control condition	Q_9	relative DES score	0.9355673	8.527429e-02
control condition	Q_{10}	relative DES score	0.9590755	3.314385e-01

Results are grouped by the two factors **treatment** (three levels: *global filter*, *peripheral filter*, and *control condition*) and **time** (11 levels: Q_0–Q_{11}).

Table 17. Shapiro–Wilk Normality Test Results for the One-factor Analysis of the Pre and Post-condition Questionnaires

	M_{all}		M_{ex}		M_{in}		M_{vr}	
	Test statistic	p-value	Test statistic	p-value	Test statistic	p-value	Test statistic	p-value
global filter	0.912	0.0222	0.748	0.0000147	0.789	0.0000698	0.907	0.0169
peripheral filter	0.894	0.00846	0.748	0.0000147	0.907	0.0172	0.797	0.0000948
no filter	0.770	0.0000327	0.634	0.000000375	0.845	0.000748	0.882	0.00450
	$M_{visualappeal}$		$M_{enjoyment}$		$M_{presence}$		$M_{obtrusiveness}$	
	Test statistic	p-value	Test statistic	p-value	Test statistic	p-value	Test statistic	p-value
global filter	0.933	0.0748	0.955	0.268	0.985	0.955	0.980	0.848
peripheral filter	0.896	0.00918	0.906	0.0155	0.955	0.263	0.983	0.906
no filter	0.907	0.0167	0.871	0.00256	0.971	0.595	0.971	0.605

Results are grouped by the two factors **treatment** (global filter, peripheral filter, and no filter).

C APPLYING EYE EXERCISES OF SHORT DURATION AND HIGH FREQUENCY TO ADDRESS DES IN VR-HMDS (USER STUDY 2)

C.1 Questionnaires

C.1.1 Post-condition Questionnaire. Questions that were presented on a 7-point Likert scale, reaching from *strongly disagree* to *strongly agree*.

Question 1: It was easy to perform the eye exercises.
Question 2: The eye exercises increased my sensation eye strain.
Question 3: The eye exercises reduced my sensation of eye strain.
Question 4: The eye exercises did not make a difference in my sensation of eye strain.
Question 5: It annoyed me doing the eye exercises.
Question 6: Doing the eye exercises was fun.
Question 7: I enjoyed doing the eye exercises.

Questions on different scales:

Question 1: Please rate the eye exercises in terms of effectiveness from -3 to 3 (-3 negative effect, 0 neutral, 3 positive effect).
Question 2: How many hours did you spent in front of digital screens today?
Question 3: In contrast to the beginning of the VR application, do your eyes feel more strained? (yes/no/don't know)
Question 4: In contrast to the beginning of the VR application, do your eyes feel less strained? (yes/no/don't know)

C.1.2 Final Questionnaire.

Question 1: Please select the three eye exercises that were most effective.
Question 2: Please select the three eye exercises that were least effective.
Question 3: Please select the three eye exercises that were most fun.
Question 4: Please select the three eye exercises that were least fun.
Question 5: Please choose your three most favourite eye exercises. Please explain.
Question 6: Please chose your three least favourite eye exercises. Please explain.
Question 7: When eye strain in VR occurs, please rate based on agreement:
- I would prefer to continue with the VR experience
- I would prefer to do eye exercises on a regular basis
- I would prefer to continue with eye strain rather than interrupting the experience doing eye exercises

C.2 Statistical Results

C.2.1 Shapiro–Wilk Normality Tests.

Table 18. Shapiro–Wilk Normality Test Results for the Two-factor Analysis of the Within-condition Questions

Condition	Time	Variable	Test statistic	p-value
close	Q_1	M_{all}	0.5366620	1.305370e-07
close	Q_2	M_{all}	0.6629938	3.390322e-06
close	Q_3	M_{all}	0.6152124	9.174424e-07
close	Q_4	M_{all}	0.8392413	1.389311e-03
control	Q_1	M_{all}	0.7558452	6.069449e-05
control	Q_2	M_{all}	0.7998072	2.924472e-04
control	Q_3	M_{all}	0.7955959	2.497989e-04
control	Q_4	M_{all}	0.8132912	4.897568e-04
E1	Q_1	M_{all}	0.6533823	2.584627e-06
E1	Q_2	M_{all}	0.5090627	6.912784e-08
E1	Q_3	M_{all}	0.4566234	2.194013e-08
E1	Q_4	M_{all}	0.5108246	7.194137e-08
E2	Q_1	M_{all}	0.8485818	2.057737e-03
E2	Q_2	M_{all}	0.8310987	9.941677e-04
E2	Q_3	M_{all}	0.8013168	3.095669e-04
E2	Q_4	M_{all}	0.7943783	2.387403e-04
E3	Q_1	M_{all}	0.8173187	5.732173e-04
E3	Q_2	M_{all}	0.8553097	2.747372e-03
E3	Q_3	M_{all}	0.8141412	5.062277e-04
E3	Q_4	M_{all}	0.7804853	1.437048e-04
E4	Q_1	M_{all}	0.8216416	6.798971e-04
E4	Q_2	M_{all}	0.8328750	1.068820e-03
E4	Q_3	M_{all}	0.8342368	1.130078e-03
E4	Q_4	M_{all}	0.8497989	2.167361e-03
E5	Q_1	M_{all}	0.7731509	1.106359e-04
E5	Q_2	M_{all}	0.6714262	4.317706e-06
E5	Q_3	M_{all}	0.6024954	6.587206e-07
E5	Q_4	M_{all}	0.8112400	4.523111e-04
E6	Q_1	M_{all}	0.8305994	9.741997e-04
E6	Q_2	M_{all}	0.8719499	5.745655e-03
E6	Q_3	M_{all}	0.8608029	3.492235e-03
E6	Q_4	M_{all}	0.8516709	2.348253e-03
E7	Q_1	M_{all}	0.9035082	2.556186e-02
E7	Q_2	M_{all}	0.8920250	1.463969e-02
E7	Q_3	M_{all}	0.8957247	1.748893e-02
E7	Q_4	M_{all}	0.7681812	9.289618e-05
E8	Q_1	M_{all}	0.7384469	3.393164e-05
E8	Q_2	M_{all}	0.8768775	7.194960e-03
E8	Q_3	M_{all}	0.9001144	2.164392e-02
E8	Q_4	M_{all}	0.6558879	2.772880e-06

(Continued)

Table 18. Continued

Condition	Time	Variable	Test statistic	p-value
close	Q_1	M_{ex}	0.4455664	1.738434e-08
close	Q_2	M_{ex}	0.6253897	1.201729e-06
close	Q_3	M_{ex}	0.7410802	3.700329e-05
close	Q_4	M_{ex}	0.7749532	1.179318e-04
control	Q_1	M_{ex}	0.7227910	2.046289e-05
control	Q_2	M_{ex}	0.7914034	2.138538e-04
control	Q_3	M_{ex}	0.7340969	2.943593e-05
control	Q_4	M_{ex}	0.8149699	5.228550e-04
E1	Q_1	M_{ex}	0.7629840	7.753770e-05
E1	Q_2	M_{ex}	0.6300218	1.360806e-06
E1	Q_3	M_{ex}	0.5218467	9.252906e-08
E1	Q_4	M_{ex}	0.5626661	2.429397e-07
E2	Q_1	M_{ex}	0.7975829	2.690319e-04
E2	Q_2	M_{ex}	0.7454841	4.281956e-05
E2	Q_3	M_{ex}	0.7746016	1.164694e-04
E2	Q_4	M_{ex}	0.7640292	8.039425e-05
E3	Q_1	M_{ex}	0.7788844	1.356813e-04
E3	Q_2	M_{ex}	0.6713672	4.310344e-06
E3	Q_3	M_{ex}	0.7875990	1.859782e-04
E3	Q_4	M_{ex}	0.7621266	7.527516e-05
E4	Q_1	M_{ex}	0.3665612	3.578983e-09
E4	Q_2	M_{ex}	0.5583411	2.187506e-07
E4	Q_3	M_{ex}	0.7567144	6.251904e-05
E4	Q_4	M_{ex}	0.7620464	7.506714e-05
E5	Q_1	M_{ex}	0.6541242	2.638918e-06
E5	Q_2	M_{ex}	0.7332194	2.860826e-05
E5	Q_3	M_{ex}	0.6202860	1.049018e-06
E5	Q_4	M_{ex}	0.7563015	6.164520e-05
E6	Q_1	M_{ex}	0.5905915	4.858900e-07
E6	Q_2	M_{ex}	0.6187903	1.008268e-06
E6	Q_3	M_{ex}	0.7299267	2.571659e-05
E6	Q_4	M_{ex}	0.7850710	1.696107e-04
E7	Q_1	M_{ex}	0.6488465	2.277468e-06
E7	Q_2	M_{ex}	0.8072595	3.880597e-04
E7	Q_3	M_{ex}	0.7908678	2.096744e-04
E7	Q_4	M_{ex}	0.8217183	6.819684e-04
E8	Q_1	M_{ex}	0.6865509	6.723300e-06
E8	Q_2	M_{ex}	0.7728579	1.094959e-04
E8	Q_3	M_{ex}	0.8647607	4.160376e-03
E8	Q_4	M_{ex}	0.6007219	6.293013e-0

(Continued)

Table 18. Continued

Condition	Time	Variable	Test statistic	p-value
close	Q_1	M_{in}	0.6815645	5.802249e-06
close	Q_2	M_{in}	0.7013252	1.048692e-05
close	Q_3	M_{in}	0.5552336	2.029521e-07
close	Q_4	M_{in}	0.8156442	5.368143e-04
control	Q_1	M_{in}	0.7771989	1.277461e-04
control	Q_2	M_{in}	0.8335928	1.100656e-03
control	Q_3	M_{in}	0.7551203	5.921616e-05
control	Q_4	M_{in}	0.7969578	2.628167e-04
E1	Q_1	M_{in}	0.7423061	3.853286e-05
E1	Q_2	M_{in}	0.6285641	1.308465e-06
E1	Q_3	M_{in}	0.4584390	2.280155e-08
E1	Q_4	M_{in}	0.5640406	2.512071e-07
E2	Q_1	M_{in}	0.7465086	4.430742e-05
E2	Q_2	M_{in}	0.9060594	2.899299e-02
E2	Q_3	M_{in}	0.8473787	1.955144e-03
E2	Q_4	M_{in}	0.8872057	1.164212e-02
E3	Q_1	M_{in}	0.6164702	9.483417e-07
E3	Q_2	M_{in}	0.5778125	3.525779e-07
E3	Q_3	M_{in}	0.4874247	4.266141e-08
E3	Q_4	M_{in}	0.6870450	6.822687e-06
E4	Q_1	M_{in}	0.7146576	1.583223e-05
E4	Q_2	M_{in}	0.7034021	1.117424e-05
E4	Q_3	M_{in}	0.7855015	1.722859e-04
E4	Q_4	M_{in}	0.8229078	7.150014e-04
E5	Q_1	M_{in}	0.7723697	1.076241e-04
E5	Q_2	M_{in}	0.6404356	1.805776e-06
E5	Q_3	M_{in}	0.7620841	7.516463e-05
E5	Q_4	M_{in}	0.8714519	5.617445e-03
E6	Q_1	M_{in}	0.5966090	5.663043e-07
E6	Q_2	M_{in}	0.8582143	3.117552e-03
E6	Q_3	M_{in}	0.7976116	2.693210e-04
E6	Q_4	M_{in}	0.9036317	2.571776e-02
E7	Q_1	M_{in}	0.8484661	2.047625e-03
E7	Q_2	M_{in}	0.8625763	3.776324e-03
E7	Q_3	M_{in}	0.8407682	1.480471e-03
E7	Q_4	M_{in}	0.7796912	1.396633e-04
E8	Q_1	M_{in}	0.6668431	3.784301e-06
E8	Q_2	M_{in}	0.7372951	3.267433e-05
E8	Q_3	M_{in}	0.8024093	3.226196e-04
E8	Q_4	M_{in}	0.6497728	2.336894e-06

(Continued)

Table 18. Continued

Condition	Time	Variable	Test statistic	p-value
close	Q_1	M_{vr}	0.5790737	3.638184e-07
close	Q_2	M_{vr}	0.6932112	8.202823e-06
close	Q_3	M_{vr}	0.5679589	2.764540e-07
close	Q_4	M_{vr}	0.6474906	2.193356e-06
control	Q_1	M_{vr}	0.7908429	2.094818e-04
control	Q_2	M_{vr}	0.7299978	2.577565e-05
control	Q_3	M_{vr}	0.8289028	9.094674e-04
control	Q_4	M_{vr}	0.7958954	2.526016e-04
E1	Q_1	M_{vr}	0.4868225	4.210010e-08
E1	Q_2	M_{vr}	0.5598357	2.268078e-07
E1	Q_3	M_{vr}	0.5613842	2.354878e-07
E1	Q_4	M_{vr}	0.5651268	2.579515e-07
E2	Q_1	M_{vr}	0.8380060	1.319933e-03
E2	Q_2	M_{vr}	0.7837355	1.615884e-04
E2	Q_3	M_{vr}	0.7227702	2.044935e-05
E2	Q_4	M_{vr}	0.8133292	4.904811e-04
E3	Q_1	M_{vr}	0.7406385	3.646805e-05
E3	Q_2	M_{vr}	0.7997221	2.915125e-04
E3	Q_3	M_{vr}	0.8987624	2.026363e-02
E3	Q_4	M_{vr}	0.7871951	1.832537e-04
E4	Q_1	M_{vr}	0.7793046	1.377398e-04
E4	Q_2	M_{vr}	0.8430472	1.628550e-03
E4	Q_3	M_{vr}	0.8124341	4.737233e-04
E4	Q_4	M_{vr}	0.7480492	4.664896e-05
E5	Q_1	M_{vr}	0.6879244	7.003416e-06
E5	Q_2	M_{vr}	0.5506738	1.819193e-07
E5	Q_3	M_{vr}	0.6021446	6.527874e-07
E5	Q_4	M_{vr}	0.7589371	6.745523e-05
E6	Q_1	M_{vr}	0.7331437	2.853812e-05
E6	Q_2	M_{vr}	0.8413078	1.514189e-03
E6	Q_3	M_{vr}	0.8282003	8.840171e-04
E6	Q_4	M_{vr}	0.7794715	1.385669e-04
E7	Q_1	M_{vr}	0.6730700	4.527981e-06
E7	Q_2	M_{vr}	0.7707981	1.018261e-04
E7	Q_3	M_{vr}	0.7793785	1.381052e-04
E7	Q_4	M_{vr}	0.7909250	2.101161e-04
E8	Q_1	M_{vr}	0.7132922	1.517084e-05
E8	Q_2	M_{vr}	0.7913751	2.136304e-04
E8	Q_3	M_{vr}	0.8166345	5.580345e-04
E8	Q_4	M_{vr}	0.8208905	6.599414e-04

(Continued)

Table 18. Continued

Condition	Time	Variable	Test statistic	p-value
close	Q_0	$M_{straining}$	0.6540880	2.636237e-06
close	Q_1	$M_{straining}$	0.7353989	3.071106e-05
close	Q_2	$M_{straining}$	0.7220321	1.997537e-05
close	Q_3	$M_{straining}$	0.7790085	1.362858e-04
close	Q_4	$M_{straining}$	0.7913567	2.134856e-04
control	Q_0	$M_{straining}$	0.7459606	4.350484e-05
control	Q_1	$M_{straining}$	0.8128401	4.812478e-04
control	Q_2	$M_{straining}$	0.9018124	2.351874e-02
control	Q_3	$M_{straining}$	0.7966260	2.595806e-04
control	Q_4	$M_{straining}$	0.8063984	3.754817e-04
E1	Q_0	$M_{straining}$	0.6469011	2.157816e-06
E1	Q_1	$M_{straining}$	0.9064731	2.959326e-02
E1	Q_2	$M_{straining}$	0.8843135	1.016070e-02
E1	Q_3	$M_{straining}$	0.9016104	2.328712e-02
E1	Q_4	$M_{straining}$	0.8511754	2.298863e-03
E2	Q_0	$M_{straining}$	0.6899064	7.429662e-06
E2	Q_1	$M_{straining}$	0.8736064	6.194884e-03
E2	Q_2	$M_{straining}$	0.8850680	1.052690e-02
E2	Q_3	$M_{straining}$	0.9210642	6.166245e-02
E2	Q_4	$M_{straining}$	0.9178592	5.238898e-02
E3	Q_0	$M_{straining}$	0.7267700	2.323452e-05
E3	Q_1	$M_{straining}$	0.9164340	4.874033e-02
E3	Q_2	$M_{straining}$	0.8861276	1.106487e-02
E3	Q_3	$M_{straining}$	0.8483457	2.037162e-03
E3	Q_4	$M_{straining}$	0.9309462	1.023528e-01
E4	Q_0	$M_{straining}$	0.6979161	9.454239e-06
E4	Q_1	$M_{straining}$	0.8993189	2.082023e-02
E4	Q_2	$M_{straining}$	0.9032769	2.527262e-02
E4	Q_3	$M_{straining}$	0.9411220	1.727869e-01
E4	Q_4	$M_{straining}$	0.8864831	1.125180e-02
E5	Q_0	$M_{straining}$	0.6720602	4.397549e-06
E5	Q_1	$M_{straining}$	0.8833222	9.699896e-03
E5	Q_2	$M_{straining}$	0.9160497	4.780223e-02
E5	Q_3	$M_{straining}$	0.9333082	1.155964e-01
E5	Q_4	$M_{straining}$	0.9460568	2.222179e-01
E6	Q_0	$M_{straining}$	0.7502036	5.014669e-05
E6	Q_1	$M_{straining}$	0.8462966	1.867512e-03
E6	Q_2	$M_{straining}$	0.8783099	7.685457e-03
E6	Q_3	$M_{straining}$	0.8947899	1.671781e-02
E6	Q_4	$M_{straining}$	0.8812567	8.809297e-03
E7	Q_0	$M_{straining}$	0.7311109	2.671908e-05
E7	Q_1	$M_{straining}$	0.8162564	5.498292e-04
E7	Q_2	$M_{straining}$	0.8104457	4.386390e-04
E7	Q_3	$M_{straining}$	0.8873096	1.169944e-02
E7	Q_4	$M_{straining}$	0.9154279	4.632372e-02
E8	Q_0	$M_{straining}$	0.6056527	7.147571e-07
E8	Q_1	$M_{straining}$	0.8802750	8.416748e-03
E8	Q_2	$M_{straining}$	0.8852559	1.062021e-02
E8	Q_3	$M_{straining}$	0.8867690	1.140453e-02
E8	Q_4	$M_{straining}$	0.8739377	6.289126e-03

(Continued)

Table 18. Continued

Condition	Time	Variable	Test statistic	p-value
close	Q_0	$M_{relieving}$	0.6875327	6.922302e-06
close	Q_1	$M_{relieving}$	0.8492425	2.116497e-03
close	Q_2	$M_{relieving}$	0.8492425	2.116497e-03
close	Q_3	$M_{relieving}$	0.8372733	1.280527e-03
close	Q_4	$M_{relieving}$	0.8629129	3.832970e-03
control	Q_0	$M_{relieving}$	0.7010046	1.038489e-05
control	Q_1	$M_{relieving}$	0.6684531	3.963245e-06
control	Q_2	$M_{relieving}$	0.7253682	2.221512e-05
control	Q_3	$M_{relieving}$	0.7243865	2.152957e-05
control	Q_4	$M_{relieving}$	0.7429266	3.933246e-05
E1	Q_0	$M_{relieving}$	0.6419372	1.881721e-06
E1	Q_1	$M_{relieving}$	0.9096204	3.460774e-02
E1	Q_2	$M_{relieving}$	0.9145852	4.439551e-02
E1	Q_3	$M_{relieving}$	0.8959067	1.764333e-02
E1	Q_4	$M_{relieving}$	0.9147817	4.483743e-02
E2	Q_0	$M_{relieving}$	0.6701140	4.157274e-06
E2	Q_1	$M_{relieving}$	0.9165539	4.903696e-02
E2	Q_2	$M_{relieving}$	0.9079311	3.181464e-02
E2	Q_3	$M_{relieving}$	0.9426013	1.863772e-01
E2	Q_4	$M_{relieving}$	0.9405937	1.681700e-01
E3	Q_0	$M_{relieving}$	0.7736597	1.126457e-04
E3	Q_1	$M_{relieving}$	0.9518960	2.976644e-01
E3	Q_2	$M_{relieving}$	0.9206524	6.038165e-02
E3	Q_3	$M_{relieving}$	0.9147669	4.480401e-02
E3	Q_4	$M_{relieving}$	0.9023291	2.412247e-02
E4	Q_0	$M_{relieving}$	0.7977144	2.703591e-04
E4	Q_1	$M_{relieving}$	0.9489771	2.574456e-01
E4	Q_2	$M_{relieving}$	0.9456168	2.173165e-01
E4	Q_3	$M_{relieving}$	0.9366869	1.375733e-01
E4	Q_4	$M_{relieving}$	0.9376458	1.445338e-01
E5	Q_0	$M_{relieving}$	0.6778202	5.199069e-06
E5	Q_1	$M_{relieving}$	0.8863969	1.120616e-02
E5	Q_2	$M_{relieving}$	0.9200092	5.843617e-02
E5	Q_3	$M_{relieving}$	0.9188243	5.501939e-02
E5	Q_4	$M_{relieving}$	0.9343967	1.222645e-01
E6	Q_0	$M_{relieving}$	0.7573934	6.398491e-05
E6	Q_1	$M_{relieving}$	0.9164185	4.870198e-02
E6	Q_2	$M_{relieving}$	0.9018053	2.351057e-02
E6	Q_3	$M_{relieving}$	0.9101135	3.547042e-02
E6	Q_4	$M_{relieving}$	0.9126771	4.033103e-02
E7	Q_0	$M_{relieving}$	0.7350432	3.035705e-05
E7	Q_1	$M_{relieving}$	0.9008968	2.248750e-02
E7	Q_2	$M_{relieving}$	0.8921396	1.472018e-02
E7	Q_3	$M_{relieving}$	0.8851846	1.058470e-02
E7	Q_4	$M_{relieving}$	0.9355316	1.296263e-01
E8	Q_0	$M_{relieving}$	0.5606132	2.311231e-07
E8	Q_1	$M_{relieving}$	0.9246627	7.411256e-02
E8	Q_2	$M_{relieving}$	0.9018707	2.358614e-02
E8	Q_3	$M_{relieving}$	0.8786631	7.811777e-03
E8	Q_4	$M_{relieving}$	0.8599922	3.370006e-03

(Continued)

Table 18. Continued

Condition	Time	Variable	Test statistic	p-value
close	Q_0	M_{tired}	0.8440042	1.695351e-03
close	Q_1	M_{tired}	0.8874767	1.179217e-02
close	Q_2	M_{tired}	0.8874283	1.176526e-02
close	Q_3	M_{tired}	0.8918183	1.449571e-02
close	Q_4	M_{tired}	0.8773012	7.336502e-03
control	Q_0	M_{tired}	0.8494453	2.134887e-03
control	Q_1	M_{tired}	0.8542744	2.626953e-03
control	Q_2	M_{tired}	0.8063455	3.747242e-04
control	Q_3	M_{tired}	0.8190826	6.144217e-04
control	Q_4	M_{tired}	0.8597490	3.334218e-03
E1	Q_0	M_{tired}	0.8865341	1.127889e-02
E1	Q_1	M_{tired}	0.8960417	1.775882e-02
E1	Q_2	M_{tired}	0.8592447	3.261317e-03
E1	Q_3	M_{tired}	0.8873096	1.169944e-02
E1	Q_4	M_{tired}	0.8239170	7.443604e-04
E2	Q_0	M_{tired}	0.8306967	9.780558e-04
E2	Q_1	M_{tired}	0.9008326	2.241707e-02
E2	Q_2	M_{tired}	0.8913311	1.416214e-02
E2	Q_3	M_{tired}	0.8948475	1.676430e-02
E2	Q_4	M_{tired}	0.9094173	3.425878e-02
E3	Q_0	M_{tired}	0.8325357	1.054114e-03
E3	Q_1	M_{tired}	0.8686434	4.948971e-03
E3	Q_2	M_{tired}	0.8765170	7.076791e-03
E3	Q_3	M_{tired}	0.8678603	4.778002e-03
E3	Q_4	M_{tired}	0.8518336	2.364716e-03
E4	Q_0	M_{tired}	0.8400337	1.435851e-03
E4	Q_1	M_{tired}	0.8446712	1.743632e-03
E4	Q_2	M_{tired}	0.9056175	2.836584e-02
E4	Q_3	M_{tired}	0.8833532	9.713971e-03
E4	Q_4	M_{tired}	0.9167193	4.944911e-02
E5	Q_0	M_{tired}	0.9132395	4.148709e-02
E5	Q_1	M_{tired}	0.8879724	1.207194e-02
E5	Q_2	M_{tired}	0.8879432	1.205527e-02
E5	Q_3	M_{tired}	0.8833135	9.695960e-03
E5	Q_4	M_{tired}	0.8556094	2.783310e-03
E6	Q_0	M_{tired}	0.9133571	4.173311e-02
E6	Q_1	M_{tired}	0.9311324	1.033389e-01
E6	Q_2	M_{tired}	0.9039662	2.614497e-02
E6	Q_3	M_{tired}	0.8948891	1.679796e-02
E6	Q_4	M_{tired}	0.9346212	1.236870e-01
E7	Q_0	M_{tired}	0.8169661	5.653407e-04
E7	Q_1	M_{tired}	0.8135864	4.954120e-04
E7	Q_2	M_{tired}	0.7164999	1.677358e-05
E7	Q_3	M_{tired}	0.7525829	5.433581e-05
E7	Q_4	M_{tired}	0.8473953	1.956528e-03
E8	Q_0	M_{tired}	0.8702498	5.320300e-03
E8	Q_1	M_{tired}	0.9097394	3.481389e-02
E8	Q_2	M_{tired}	0.8923015	1.483470e-02
E8	Q_3	M_{tired}	0.9131566	4.131460e-02
E8	Q_4	M_{tired}	0.8572293	2.986406e-03

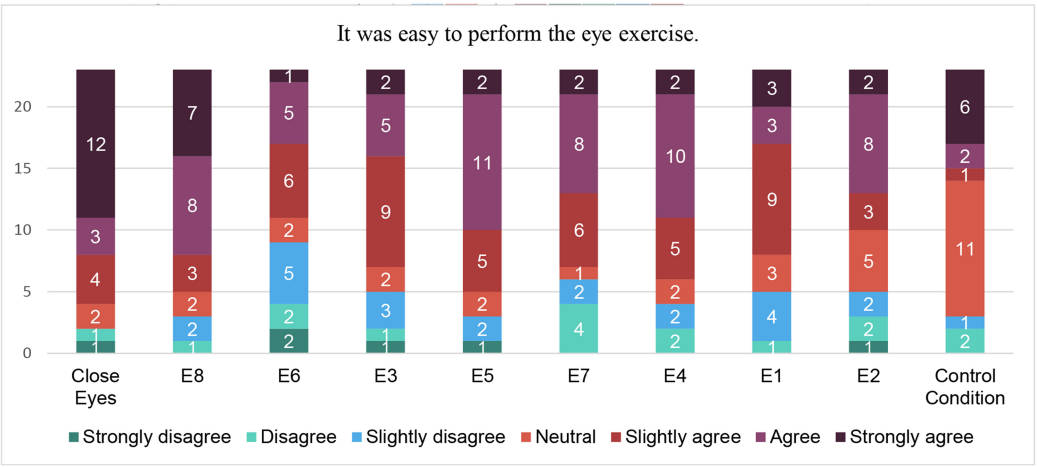
(Continued)

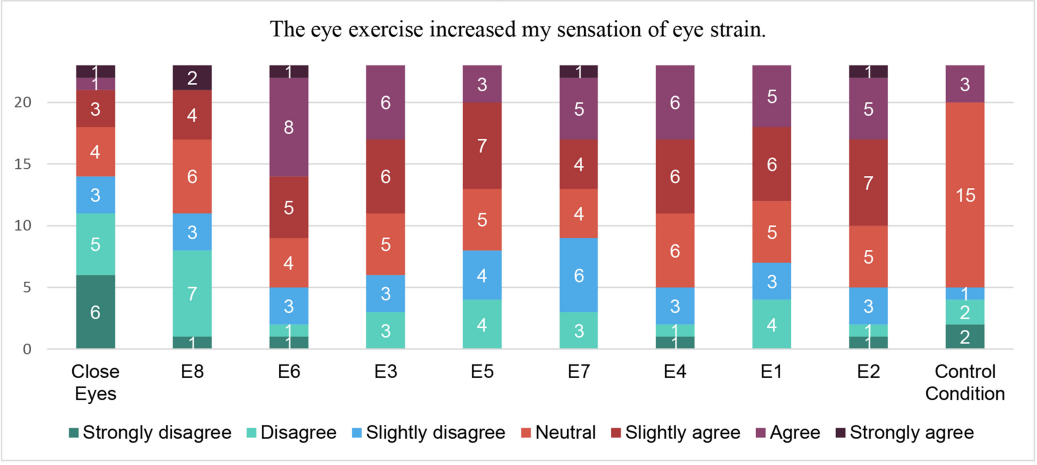
Table 18. Continued

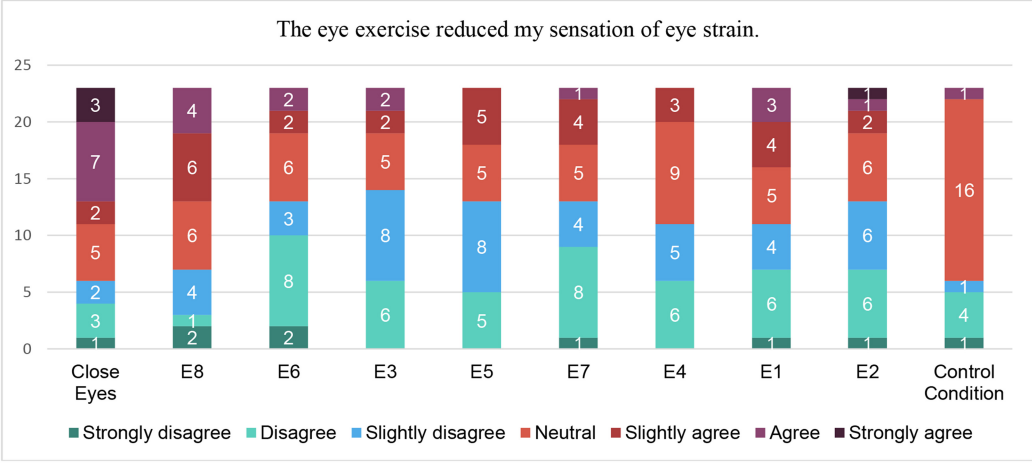
Condition	Time	Variable	Test statistic	p-value
close	Q_0	$M_{relaxed}$	0.8675625	0.0047146465
close	Q_1	$M_{relaxed}$	0.8648385	0.0041747923
close	Q_2	$M_{relaxed}$	0.8990887	0.0205880259
close	Q_3	$M_{relaxed}$	0.9158892	0.0474160210
close	Q_4	$M_{relaxed}$	0.8696677	0.0051824458
control	Q_0	$M_{relaxed}$	0.8753778	0.0067167441
control	Q_1	$M_{relaxed}$	0.8392699	0.0013909650
control	Q_2	$M_{relaxed}$	0.8173717	0.0005744131
control	Q_3	$M_{relaxed}$	0.8744338	0.0064330739
control	Q_4	$M_{relaxed}$	0.9142827	0.0437237153
E1	Q_0	$M_{relaxed}$	0.8639586	0.0040147535
E1	Q_1	$M_{relaxed}$	0.8579421	0.0030807201
E1	Q_2	$M_{relaxed}$	0.8230195	0.0007181879
E1	Q_3	$M_{relaxed}$	0.8548727	0.0026958379
E1	Q_4	$M_{relaxed}$	0.8278288	0.0008708653
E2	Q_0	$M_{relaxed}$	0.9198573	0.0579862505
E2	Q_1	$M_{relaxed}$	0.8642951	0.0040751743
E2	Q_2	$M_{relaxed}$	0.8952386	0.0170834113
E2	Q_3	$M_{relaxed}$	0.9232458	0.0689282494
E2	Q_4	$M_{relaxed}$	0.8860146	0.0110061112
E3	Q_0	$M_{relaxed}$	0.8583627	0.0031378408
E3	Q_1	$M_{relaxed}$	0.8549176	0.0027010940
E3	Q_2	$M_{relaxed}$	0.9530561	0.3151322797
E3	Q_3	$M_{relaxed}$	0.9119996	0.0389826257
E3	Q_4	$M_{relaxed}$	0.9215322	0.0631518852
E4	Q_0	$M_{relaxed}$	0.8405252	0.0014655440
E4	Q_1	$M_{relaxed}$	0.8653591	0.0042726660
E4	Q_2	$M_{relaxed}$	0.9071353	0.0305813771
E4	Q_3	$M_{relaxed}$	0.9052372	0.0278374366
E4	Q_4	$M_{relaxed}$	0.8900313	0.0133112504
E5	Q_0	$M_{relaxed}$	0.9409284	0.1710814580
E5	Q_1	$M_{relaxed}$	0.8950910	0.0169621835
E5	Q_2	$M_{relaxed}$	0.8697641	0.0052050061
E5	Q_3	$M_{relaxed}$	0.9112749	0.0375921584
E5	Q_4	$M_{relaxed}$	0.9207204	0.0605911945
E6	Q_0	$M_{relaxed}$	0.8799892	0.0083059635
E6	Q_1	$M_{relaxed}$	0.9243535	0.0729479094
E6	Q_2	$M_{relaxed}$	0.9427584	0.1878796124
E6	Q_3	$M_{relaxed}$	0.9132271	0.0414612069
E6	Q_4	$M_{relaxed}$	0.9254451	0.0771443697
E7	Q_0	$M_{relaxed}$	0.8616722	0.0036285511
E7	Q_1	$M_{relaxed}$	0.8591281	0.0032446976
E7	Q_2	$M_{relaxed}$	0.7959647	0.0002532548
E7	Q_3	$M_{relaxed}$	0.8641200	0.0040436100
E7	Q_4	$M_{relaxed}$	0.8984719	0.0199792827
E8	Q_0	$M_{relaxed}$	0.9101135	0.0354704216
E8	Q_1	$M_{relaxed}$	0.9350049	0.1261564199
E8	Q_2	$M_{relaxed}$	0.9083408	0.0324697107
E8	Q_3	$M_{relaxed}$	0.8902305	0.0134380301
E8	Q_4	$M_{relaxed}$	0.9265803	0.0817705317

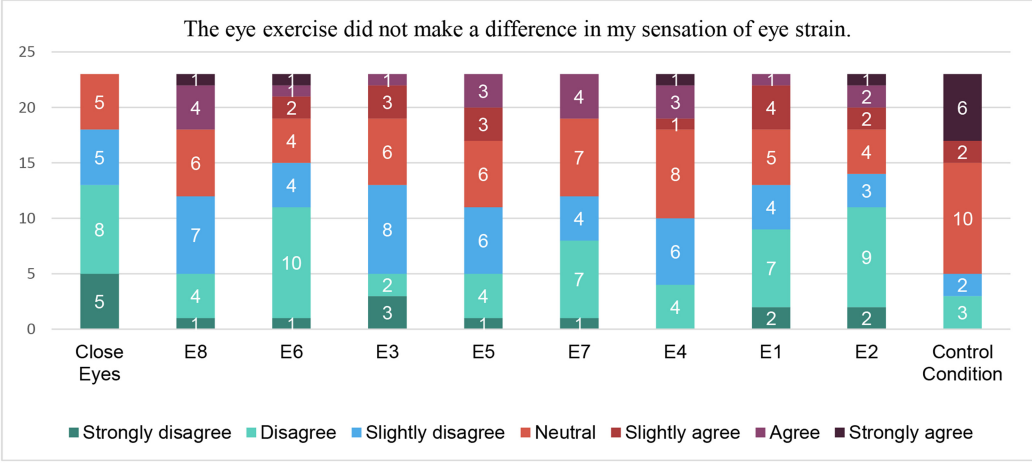
Results are grouped by the two factors **treatment** (10 levels: *closing the eyes*, *control condition*, *E1–E8*) and **time** (four levels: Q_0–Q_3).

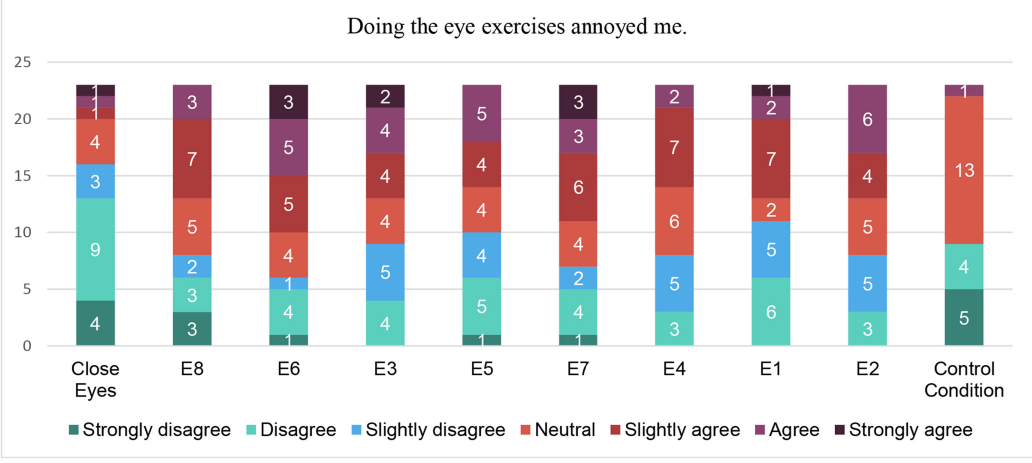
C.2.2 Post-condition Questionnaire. Results of the post-condition questionnaire of user study 2 in which participants answered 13 statements about the eye exercises.

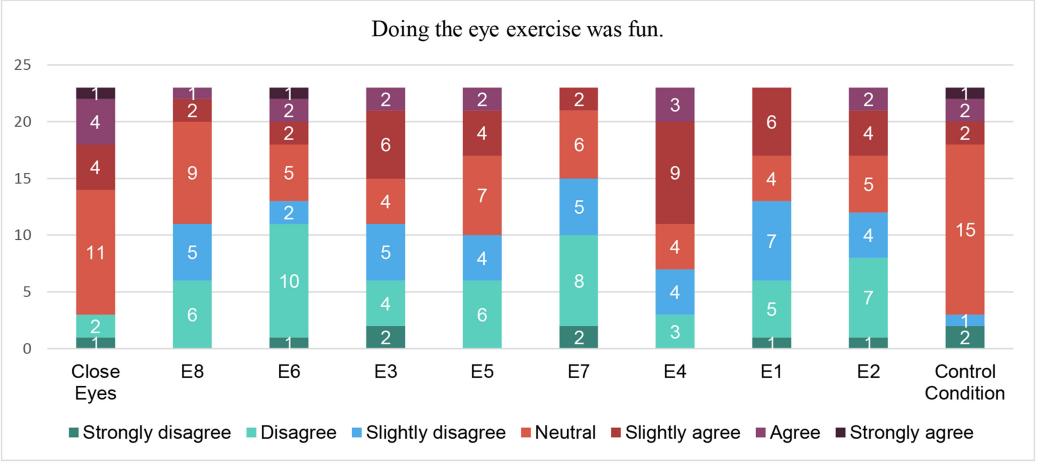


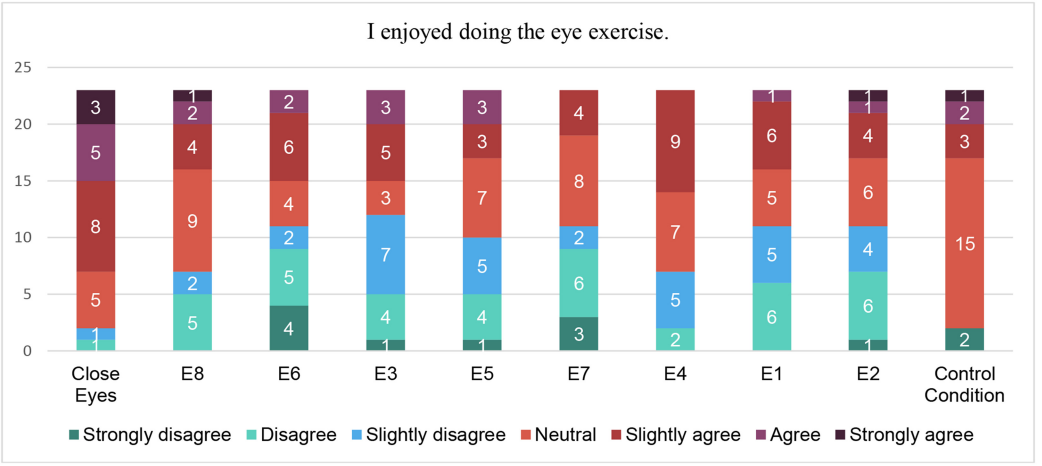


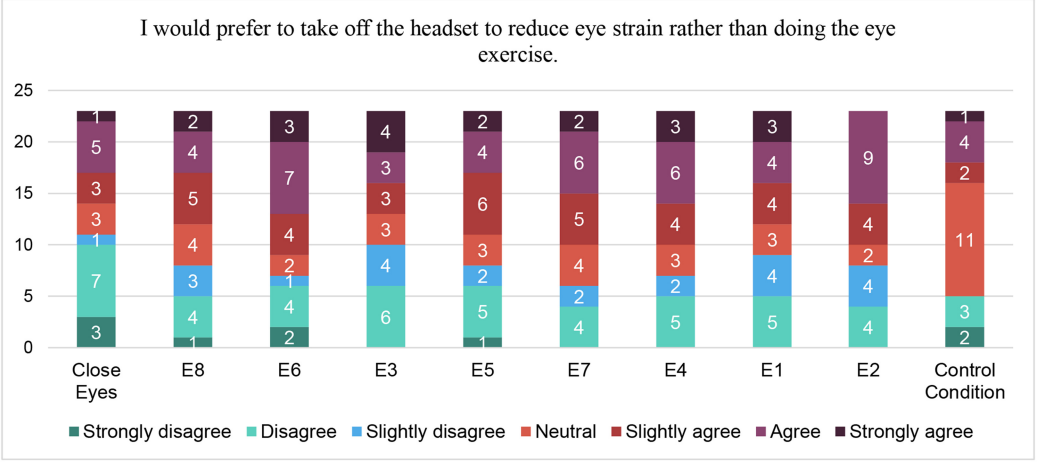

The eye exercise reduced my sensation of eye strain.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree

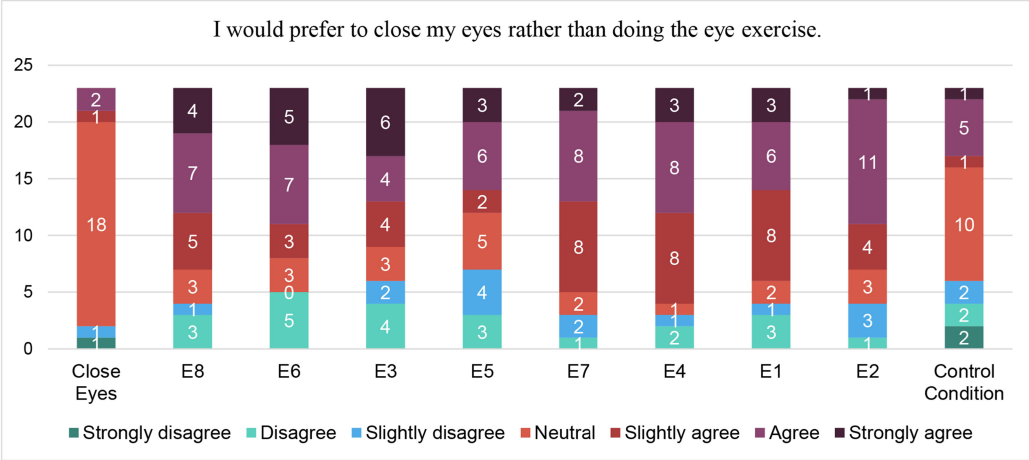

The eye exercise did not make a difference in my sensation of eye strain.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree

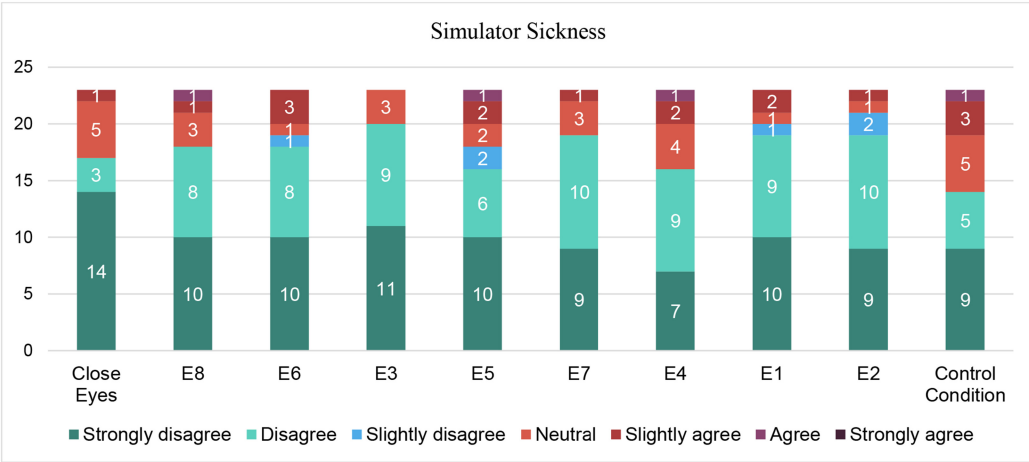

Doing the eye exercises annoyed me.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree

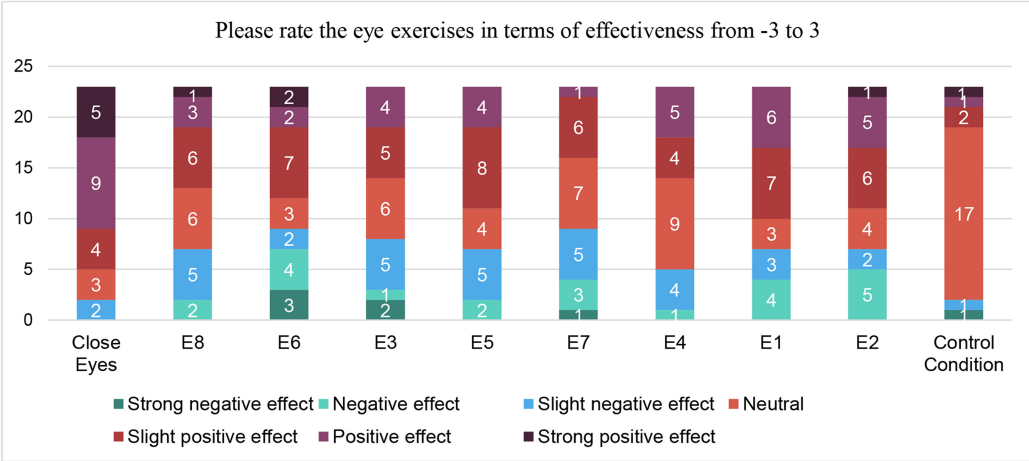

Doing the eye exercise was fun.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree

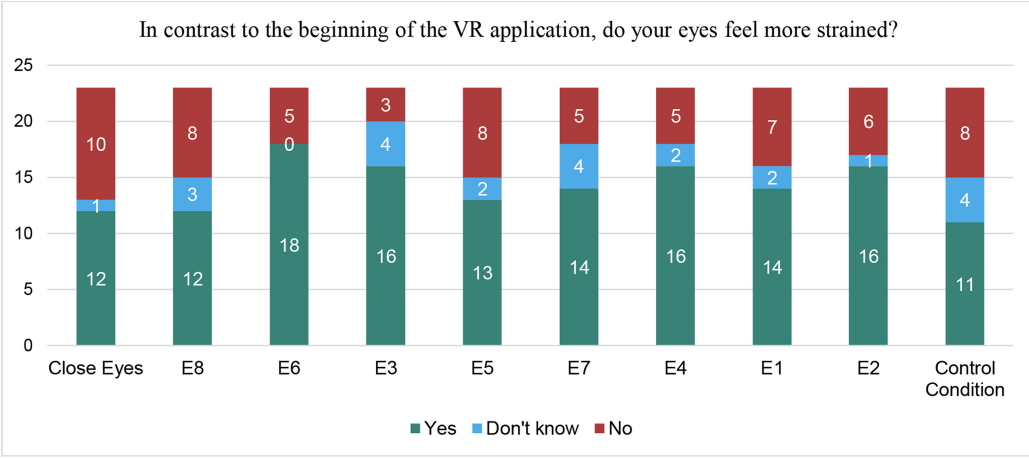

I enjoyed doing the eye exercise.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree

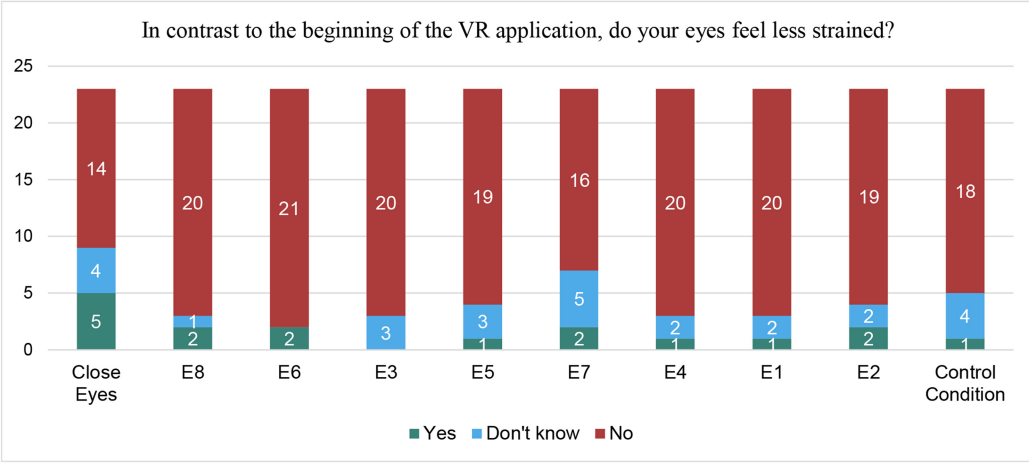

I would prefer to take off the headset to reduce eye strain rather than doing the eye exercise.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree


I would prefer to close my eyes rather than doing the eye exercise.
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree


Simulator Sickness
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strongly disagree
Disagree
Slightly disagree
Neutral
Slightly agree
Agree
Strongly agree


Please rate the eye exercises in terms of effectiveness from -3 to 3
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
Strong negative effect
Negative effect
Slight negative effect
Neutral
Slight positive effect
Positive effect
Strong positive effect


In contrast to the beginning of the VR application, do your eyes feel more strained?
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
10
1
12
8
3
12
5
0
18
3
4
16
8
2
13
5
4
14
5
2
16
7
2
14
6
1
16
8
4
11
Yes
Don't know
No


In contrast to the beginning of the VR application, do your eyes feel less strained?
25
20
15
10
5
0
Close Eyes
E8
E6
E3
E5
E7
E4
E1
E2
Control Condition
14
4
5
20
1
2
21
2
20
3
19
3
1
16
5
2
20
2
1
20
2
1
19
2
2
18
4
1
Yes
Don't know
No

D APPLYING EYE EXERCISES OF LONG DURATION AND LOW FREQUENCY TO ADDRESS DES IN VR HMDS (USER STUDY 3)

D.1 Shapiro–Wilk Normality Tests

Table 19. Shapiro-Wilk Normality Test Results for the Two-factor Analysis of the Within-condition Questions

Condition	Time	Variable	Test statistic	p-value
Close Eyes	Q_{1-0}	M_{all}	0.96557	0.6856
Close Eyes	Q_{2-0}	M_{all}	0.76605	0.0003823
Close Eyes	Q_{3-0}	M_{all}	0.85401	0.007801
Eye Exercises	Q_{1-0}	M_{all}	0.93898	0.2528
Eye Exercises	Q_{2-0}	M_{all}	0.90129	0.05134
Eye Exercises	Q_{3-0}	M_{all}	0.98438	0.9808
Control Condition	Q_{1-0}	M_{all}	0.90469	0.05919
Control Condition	Q_{2-0}	M_{all}	0.92903	0.1662
Control Condition	Q_{3-0}	M_{all}	0.85047	0.006828
Close Eyes	Q_{1-0}	M_{ex}	0.95345	0.4513
Close Eyes	Q_{2-0}	M_{ex}	0.81102	0.00166
Close Eyes	Q_{3-0}	M_{ex}	0.86758	0.01314
Eye Exercises	Q_{1-0}	M_{ex}	0.98621	0.9899
Eye Exercises	Q_{2-0}	M_{ex}	0.96676	0.7103
Eye Exercises	Q_{3-0}	M_{ex}	0.98415	0.9794
Control Condition	Q_{1-0}	M_{ex}	0.9516	0.4206
Control Condition	Q_{2-0}	M_{ex}	0.94824	0.3687
Control Condition	Q_{3-0}	M_{ex}	0.87102	0.01503
Close Eyes	Q_{1-0}	M_{in}	0.9428	0.2959
Close Eyes	Q_{2-0}	M_{in}	0.71201	7.756e-05
Close Eyes	Q_{3-0}	M_{in}	0.90919	0.07153
Eye Exercises	Q_{1-0}	M_{in}	0.93659	0.2288
Eye Exercises	Q_{2-0}	M_{in}	0.79604	0.001002
Eye Exercises	Q_{3-0}	M_{in}	0.94401	0.311
Control Condition	Q_{1-0}	M_{in}	0.91084	0.07669
Control Condition	Q_{2-0}	M_{in}	0.91873	0.1072
Control Condition	Q_{3-0}	M_{in}	0.91195	0.08041
Close Eyes	Q_{1-0}	M_{vr}	0.83956	0.004558
Close Eyes	Q_{2-0}	M_{vr}	0.93215	0.1897
Close Eyes	Q_{3-0}	M_{vr}	0.75515	0.0002734
Eye Exercises	Q_{1-0}	M_{vr}	0.92619	0.1473
Eye Exercises	Q_{2-0}	M_{vr}	0.83485	0.00384
Eye Exercises	Q_{3-0}	M_{vr}	0.88921	0.03117
Control Condition	Q_{1-0}	M_{vr}	0.86675	0.01272
Control Condition	Q_{2-0}	M_{vr}	0.90908	0.0712
Control Condition	Q_{3-0}	M_{vr}	0.88491	0.02616

Results are grouped by the two factors **treatment** (three levels: *closing the eyes*, *eye exercises*, and *control condition*) and **time** (three levels: Q_{1-0}–Q_{3-0}). We show the results for the relative symptoms means.

D.2 Friedman Tests

	M_{all}	M_{ex}	M_{in}	M_{vr}
Friedman test results	$\chi^2(2) = 4.03, p = .13$	$\chi^2(2) = 2.08, p = .35$	$\chi^2(2) = 1.74, p = .42$	$\chi^2(2) = 1.63, p = .44$

D.3 Two-factorial Non-parametric Variance Analysis

	M_{all}	M_{ex}	M_{in}	M_{vr}
Time	$\chi^2(2) = 13.65, p < .01$	$\chi^2(2) = 5.88, p < .01$	$\chi^2(2) = 17.68, p < .01$	$\chi^2(2) = 9.91, p < .01$
Condition	$\chi^2(2) = 3.41, p = .04$	$\chi^2(2) = 3.18, p = .04$	$\chi^2(2) = 1.82, p = .16$	$\chi^2(2) = 2.25, p = .11$
Time:Condition	$\chi^2(3) = 4.87, p < .01$	$\chi^2(3) = 3.86, p < .01$	$\chi^2(2) = 4.35, p < .01$	$\chi^2(3) = 3.20, p = .02$

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135, 1 (2020), 102370. DOI: https://doi.org/10.1016/j.ijhcs.2019.102370

[2] Seon Jong Ahn. 2017. Device for Treating Dry Eye Syndrome and Strengthening Eyesight. US Patent App. 15/512,735.

[3] Majed Al Zayer, Isayas B. Adhanom, Paul MacNeilage, and Eelke Folmer. 2019. The effect of field-of-view restriction on sex bias in VR sickness and spatial navigation performance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–12. DOI: https://doi.org/10.1145/3290605.3300584

[4] Shelly L. Ames, James S. Wolffsohn, and Neville A. McBrien. 2005. The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display. 82, 3 (2005), 168–176. Retrieved from https://journals.lww.com/optvissci/Fulltext/2005/03000/The_Development_of_a_Symptom_Questionnaire_for.8.aspx.

[5] Jeffrey Anshel. 2005. Computer vision syndrome. *Visual Ergonomics Handbook* (2005), 23–35. DOI: https://doi.org/10.1201/9781420032055

[6] Terry Bahill and Lawrence Stark. 1975. Overlapping saccades and glissades are produced by fatigue in the saccadic eye movement system. *Experimental Neurology* 48, 1 (1975), 95–106. DOI: https://doi.org/10.1016/0014-4886(75)90225-3

[7] Jackson Beatty and Brennis Lucero-Wagoner. 2000. The pupillary system. *Handbook of Psychophysiology* 2 (2000), 142–162.

[8] Maximino Bessa, Miguel Melo, David Narciso, Luís Barbosa, and José Vasconcelos-Raposo. 2016. Does 3D 360 video enhance user's VR experience? An evaluation study. In *Proceedings of the 17th International Conference on Human Computer Interaction*. ACM, New York, NY, Article 16, 4 pages. DOI: https://doi.org/10.1145/2998626.2998669

[9] Pradipta Biswas and Pat Langdon. 2013. A new interaction technique involving eye gaze tracker and scanning system. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*. ACM, New York, NY, 67–70. DOI: https://doi.org/10.1145/2509315.2509322

[10] James Blaha and Manish Gupta. 2014. Diplopia: A virtual reality game designed to help amblyopics. In *Proceedings of the 2014 IEEE Virtual Reality*. IEEE, New York, NY, 163–164. DOI: https://doi.org/10.1109/VR.2014.6802102

[11] James Blaha and Manish Gupta. 2017. Interactive System for Vision Assessment and Correction. US Patent 9,706,910.

[12] Clayton Blehm, Seema Vishnu, Ashbala Khattak, Shrabanee Mitra, and Richard W. Yee. 2005. Computer vision syndrome: A review. *Survey of Ophthalmology* 50, 3 (2005), 253–262. DOI: https://doi.org/10.1016/j.survophthal.2005.02.008

[13] Celeste V. Bonham and Richard D. Rallison. 1992. Optical System Which Helps Reduce Eye Strain. US Patent 5,173,724.

[14] Gunnar Borg. 1998. *Borg's Perceived Exertion and Pain Scales*. Human Kinetics, Champaign, IL. Pages: viii, 104.

[15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. DOI: https://doi.org/10.1191/1478088706qp063oa

[16] Jorge Cancela, Matteo Pastorino, Alexandros T. Tzallas, Markos G. Tsipouras, Giorgios Rigas, Maria T. Arredondo, and Dimitrios I. Fotiadis. 2014. Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. *Sensors (Basel, Switzerland)* 14, 9 (September 2014), 17235–17255. DOI: https://doi.org/10.3390/s140917235 Publisher: MDPI.

[17] Marcus Carter, Joshua Newn, Eduardo Velloso, and Frank Vetere. 2015. Remote gaze and gesture tracking on the microsoft kinect: Investigating the role of feedback. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*. ACM, New York, NY, 167–176. DOI: https://doi.org/10.1145/2838739.2838778

[18] Tuhin Chakraborty, Sayan Sarcar, and Debasis Samanta. 2014. Design and evaluation of a dwell-free eye typing technique. In *Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1573–1578. DOI: https://doi.org/10.1145/2559206.2581265

[19] Chih-Wei Chen, Tommi Määttä, Kevin Bing-Yung Wong, and Hamid Aghajan. 2012. A collaborative framework for ergonomic feedback using smart cameras. In *Proceedings of the 2012 6th International Conference on Distributed Smart Cameras*. 1–6.

[20] Siyuan Chen and Julien Epps. 2014. Using task-induced pupil diameter and blink rate to infer cognitive load. *Human–Computer Interaction* 29, 4 (2014), 390–413. DOI: https://doi.org/10.1080/07370024.2014.892428

[21] Wei Chen, Jian-Gang Chao, Xue-Wen Chen, Jin-Kun Wang, and Cheng Tan. 2015. Quantitative orientation preference and susceptibility to space motion sickness simulated in a virtual reality environment. *Brain Research Bulletin* 113, 1 (2015), 17–26. DOI: https://doi.org/10.1016/j.brainresbull.2015.01.007

[22] Hong-Ming Cheng, Shyan-Tarng Chen, Liu Hsiang-Jui, and Y. Cheng. 2014. Does blue light filter improve computer vision syndrome in patients with dry eye. *Life Science Journal* 11, 6 (2014), 612–615.

[23] Christina A. Chu, Mark Rosenfield, and Joan K. Portello. 2014. Blink patterns: Reading from a computer screen versus hard copy. *Optometry and Vision Science* 91, 3 (2014), 297–302. DOI: https://doi.org/10.1097/OPX.0000000000000157

[24] Chantal Coles-Brennan, Anna Sulley, and Graeme Young. 2019. Management of digital eye strain. *Clinical and Experimental Optometry* 102, 1 (2019), 18–29. DOI: https://doi.org/10.1111/cxo.12798

[25] Juanita D. Collier and Mark Rosenfield. 2011. Accommodation and convergence during sustained computer work. *Optometry - Journal of the American Optometric Association* 82, 7 (2011), 434–440. DOI: https://doi.org/10.1016/j.optm.2010.10.013

[26] Marvin J. Dainoff, Alan Happ, and Peter Crane. 1981. Visual fatigue and occupational stress in VDT operators. *Human Factors* 23, 4 (1981), 421–437. DOI: https://doi.org/10.1177/001872088102300405 arXiv:https://doi.org/10.1177/001872088102300405. PMID: 7275110.

[27] Hugh Davson. 1990. *Physiology of the Eye*. Macmillan International Higher Education.

[28] María del Mar Seguí, Julio Cabrero-García, Ana Crespo, José Verdú, and Elena Ronda. 2015. A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. *Journal of Clinical Epidemiology* 68, 6 (2015), 662–673. DOI: https://doi.org/10.1016/j.jclinepi.2015.01.015

[29] Artem Dementyev and Christian Holz. 2017. DualBlink: A wearable device to continuously detect, track, and actuate blinking for alleviating dry eyes and computer vision syndrome. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 1, Article 1 (March 2017), 19 pages. DOI: https://doi.org/10.1145/3053330

[30] Matjaz Divjak and Horst Bischof. 2009. Eye blink based fatigue detection for prevention of computer vision syndrome. In *Proceedings of IAPR Conference on Machine Vision Applications*. 350–353.

[31] Shauney L. Donoghue. 1983. Dealing with the video display terminal. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique* 74, 3 (1983), 179–182. Retrieved from http://www.jstor.org/stable/41987053.

[32] Clair A. Enthoven, Jan Roelof Polling, Timo Verzijden, J. Willem L. Tideman, Nora Al-Jaffar, Pauline W. Jansen, Hein Raat, Lauwerens Metz, Virginie J. M. Verhoeven, and Caroline C. W. Klaver. 2021. Smartphone use associated with refractive error in teenagers: The myopia app study. *Ophthalmology* 128, 12 (2021), 1681–1688. DOI: https://doi.org/10.1016/j.ophtha.2021.06.016

[33] Moira B. Flanagan, James G. May, and Thomas G. Dobie. 2005. Sex differences in tolerance to visually-induced motion sickness. *Aviation, Space, and Environmental Medicine* 76, 7 (2005), 642–646. Retrieved from https://www.ingentaconnect.com/content/asma/asem/2005/00000076/00000007/art00004.

[34] Sebastian Freitag, Benjamin Weyers, and Torsten W. Kuhlen. 2016. Examining rotation gain in CAVE-like virtual environments. In *Proceedings of the IEEE Transactions on Visualization and Computer Graphics*. IEEE, New York, NY, 1462–1471. DOI: https://doi.org/10.1109/TVCG.2016.2518298

[35] Julian Frommel, Katja Rogers, Julia Brich, Daniel Besserer, Leonard Bradatsch, Isabel Ortinau, Ramona Schabenberger, Valentin Riemer, Claudia Schrader, and Michael Weber. 2015. Integrated questionnaires: Maintaining presence in game environments for self-reported data acquisition. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. Association for Computing Machinery, New York, NY, 359–368. DOI: https://doi.org/10.1145/2793107.2793130

[36] Angelo Gargantini, Fabio Terzi, Matteo Zambelli, and Silvia Bonfanti. 2015. A low-cost virtual reality game for amblyopia rehabilitation. In *Proceedings of the 3rd 2015 Workshop on ICTs for Improving Patients Rehabilitation Research Techniques*. Association for Computing Machinery, New York, NY, 81–84. DOI: https://doi.org/10.1145/2838944.2838964

[37] Jan Gugenheimer, Christian Mai, Mark McGill, Julie Williamson, Frank Steinicke, and Ken Perlin. 2019. Challenges using head-mounted displays in shared and social spaces. In *Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article W19, 8 pages. DOI: https://doi.org/10.1145/3290607.3299028

[38] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, and Andreas Bulling. A Survey on Digital Eye Strain in Gaze-Based Interactive Systems. 2020. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA'20 Full Papers* (Stuttgart, Germany). Association for Computing Machinery, New York, NY. DOI: https://doi.org/10.1145/3379155.3391313

[39] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. 2021. A critical assessment of the use of SSQ as a measure of general discomfort in VR head-mounted displays. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, 12 pages. DOI: https://doi.org/10.1145/3411764.3445361

[40] Teresa Hirzle, Jan Gugenheimer, Florian Geiselhart, Andreas Bulling, and Enrico Rukzio. 2019. A design space for gaze interaction on head-mounted displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, Article 625, 12 pages. DOI: https://doi.org/10.1145/3290605.3300855

[41] Jimmy Ho, Reinhard Pointner, Huai-Chun Shih, Yu-Chih Lin, Hsuan-Yu Chen, Wei-Luan Tseng, and Mike Y. Chen. 2015. EyeProtector: Encouraging a healthy viewing distance when using smartphones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. Association for Computing Machinery, New York, NY, 77–85. DOI: https://doi.org/10.1145/2785830.2785836

[42] David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. 2008. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 8, 3 (2008), 33–33. DOI: https://doi.org/10.1167/8.3.33

[43] Anna M. Horwood, Sonia S. Toor, and Patricia M. Riddell. 2014. Change in convergence and accommodation after two weeks of eye exercises in typical young adults. *Journal of American Association for Pediatric Ophthalmology and Strabismus* 18, 2 (2014), 162–168. DOI: https://doi.org/10.1016/j.jaapos.2013.11.008

[44] Peter Alan Howarth and Howell Owen Istance. 1985. The association between visual discomfort and the use of visual display units. *Behaviour & Information Technology* 4, 2 (1985), 131–149. DOI: https://doi.org/10.1080/01449298508901794

[45] Kuan-Hong Hsieh and Xiao-Guang Li. 2008. System and Method for Avoiding Eye and Bodily Injury from Using a Display Device. US Patent 7,446,762.

[46] Takeshi Ide, Ikuko Toda, Emiko Miki, and Kazuo Tsubota. 2015. Effect of blue light–reducing eye glasses on critical flicker frequency. *The Asia-Pacific Journal of Ophthalmology* 4, 2 (2015), 80–85. Retrieved from https://journals.lww.com/apjoo/Fulltext/2015/03000/Effect_of_Blue_Light_Reducing_Eye_Glasses_on.6.aspx.

[47] Wolfgang Jaschinski, Matthias Bonacker, and Ewald Alshuth. 1996. Accommodation, convergence, pupil diameter and eye blinks at a CRT display flickering near fusion limit. *Ergonomics* 39, 1 (1996), 152–164. DOI: https://doi.org/10.1080/00140139608964441. PMID: 8851079.

[48] Christian Jerome, Richard Darnell, Brian Oakley, and Aaron Pepe. 2005. The effects of presence and time of exposure on simulator sickness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 49, 26 (2005), 2258–2262. DOI: https://doi.org/10.1177/154193120504902609

[49] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, 1151–1160. DOI: https://doi.org/10.1145/2638728.2641695

[50] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. DOI: https://doi.org/10.1207/s15327108ijap0303_3

[51] Hyun K. Kim, Jaehyun Park, Yeongcheol Choi, and Mungyeong Choe. 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics* 69, 1 (2018), 66–73. https://doi.org/10.1016/j.apergo.2017.12.016

[52] Joohwan Kim, David Kane, and Martin S. Banks. 2014. The rate of change of vergence–accommodation conflict affects visual discomfort. *Vision Research* 105, 1 (2014), 159–165. DOI: https://doi.org/10.1016/j.visres.2014.10.021

[53] Kangsoo Kim, Austin Erickson, Alexis Lambert, Gerd Bruder, and Greg Welch. 2019. Effects of dark mode on visual fatigue and acuity in optical see-through head-mounted displays. In *Proceedings of the Symposium on Spatial User Interaction*. ACM, New York, NY, Article 9, 9 pages. DOI: https://doi.org/10.1145/3357251.3357584

[54] Jeff Klingner. 2010. Fixation-aligned pupillary response averaging. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. Association for Computing Machinery, New York, NY, 275–282. DOI: https://doi.org/10.1145/1743666.1743732

[55] James F. Knight and Chris Baber. 2005. A tool to assess the comfort of wearable computers. *Human Factors* 47, 1 (2005), 77–91. DOI: https://doi.org/10.1518/0018720053653875 PMID: 15960088.

[56] Takashi Kojima, Osama M. A. Ibrahim, Tais Wakamatsu, Atsushi Tsuyama, Junko Ogawa, Yukihiro Matsumoto, Murat Dogru, and Kazuo Tsubota. 2011. The impact of contact lens wear and visual display terminal work on ocular surface and tear functions in office workers. *American Journal of Ophthalmology* 152, 6 (2011), 933–940.e2. DOI: https://doi.org/10.1016/j.ajo.2011.05.025

[57] Rotem Kronenberg and Tsvi Kuflik. 2019. Automatically adjusting computer screen. In *Proceedings of the Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*. Association for Computing Machinery, New York, NY, 51–56. DOI: https://doi.org/10.1145/3314183.3324980

[58] Junko Kuze and Kazuhiko Ukai. 2008. Subjective evaluation of visual fatigue caused by motion images. *Displays* 29, 2 (2008), 159–166. DOI: https://doi.org/10.1016/j.displa.2007.09.007. Health and Safety Aspects of Visual Displays.

[59] Marc Lambooij, Marten Fortuin, and Ingrid Heynderickx. 2009. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology* 53, 3 (2009), 030201. DOI: https://doi.org/10.2352/J.ImagingSci.Technol.2009.53.3.030201

[60] Eike Langbehn, Frank Steinicke, Markus Lappe, Gregory F. Welch, and Gerd Bruder. 2018. In the blink of an eye: Leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. *ACM Transactions on Graphics* 37, 4, Article 66 (July 2018), 11 pages. DOI: https://doi.org/10.1145/3197517.3201335

[61] Ky Loh and Sagili Chandrasekhara Reddy. 2008. Understanding and preventing computer vision syndrome. *Malaysian Family Physician: The official journal of the academy of family physicians of malaysia* 3, 3 (Dec. 2008), 128–130. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4170366/.

[62] Christian Mai, Mariam Hassib, and Rolf Königbauer. 2017. Estimating visual discomfort in head-mounted displays using electroencephalography. In *Proceedings of the IFIP Conference on Human-Computer Interaction*. Springer, 243–252. DOI: https://doi.org/10.1007/978-3-319-68059-0_15

[63] Behzad Mansouri, Khizer Khaderi, Jing Yuan, Jeffrey Pettey, and Sarah Ahmed. 2018. The visual effects associated with head-mounted displays. *International Journal of Ophthalmology and Clinical Research* 5, 2 (2018), 085. DOI: https://doi.org/10.23937/2378-346X/1410085

[64] Biljana Miljanović, Reza Dana, David A. Sullivan, and Debra A. Schaumberg. 2007. Impact of dry eye syndrome on vision-related quality of life. *American Journal of Ophthalmology* 143, 3 (2007), 409–415.e2. DOI: https://doi.org/10.1016/j.ajo.2006.11.060

[65] Chulhong Min, Euihyeok Lee, Souneil Park, and Seungwoo Kang. 2019. Tiger: Wearable glasses for the 20-20-20 rule to alleviate computer vision syndrome. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. Association for Computing Machinery, New York, NY, Article 6, 11 pages. DOI: https://doi.org/10.1145/3338286.3340117

[66] Aanand Nayyar, Utkarsh Dwivedi, Karan Ahuja, Nitendra Rajput, Seema Nagar, and Kuntal Dey. 2017. OptiDwell: Intelligent adjustment of dwell click time. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. ACM, New York, NY, 193–204. DOI: https://doi.org/10.1145/3025171.3025202

[67] Kimihiro Noguchi, Yulia R. Gel, Edgar Brunner, and Frank Konietschke. 2012. nparLD: An R software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software* 50, 12 (2012), 1–23. Retrieved from http://www.jstatsoft.org/v50/i12/.

[68] U.S. National Library of Medicine. 1999. Asthenopia. Retrieved on October 22, 2021 from https://meshb.nlm.nih.gov/record/ui?ui=D001248.

[69] Tatsiana Palavets and Mark Rosenfield. 2019. Blue-blocking filters and digital eyestrain. *Optometry and Vision Science* 96, 1 (2019), 48–54. DOI: https://doi.org/10.1097/OPX.0000000000001318

[70] Timo Partala and Veikko Surakka. 2003. Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies* 59, 1 (2003), 185–198. DOI: https://doi.org/10.1016/S1071-5819(03)00017-X

[71] Eli Peli. 1998. The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display. *Vision Research* 38, 13 (1998), 2053–2066. DOI: https://doi.org/10.1016/S0042-6989(97)00397-0

[72] Eli Peli. 1999. *Optometric and Perceptual Issues with Head-mounted Displays*. McGraw-Hill New York, New York, NY, 205–276.

[73] Joan K. Portello, Mark Rosenfield, and Christina A. Chu. 2013. Blink rate, incomplete blinks and computer vision syndrome. *Optometry and Vision Science* 90, 5 (2013), 482–487. DOI: https://doi.org/10.1097/OPX.0b013e31828f09a7

[74] Photchara Ratsamee, Yasushi Mae, Kazuto Kamiyama, Mitsuhiro Horade, Masaru Kojima, Kiyoshi Kiyokawa, Tomohiro Mashita, Yoshihiro Kuroda, Haruo Takemura, and Tatsuo Arai. 2015. Object search framework based on gaze interaction. In *Proceedings of the 2015 IEEE International Conference on Robotics and Biomimetics*. IEEE, New York, NY, 1997–2002. DOI: https://doi.org/10.1109/ROBIO.2015.7419066

[75] John Albert Rawstron, Celeste D. Burley, and Mark J. Elder. 2005. A systematic review of the applicability and efficacy of eye exercises. *Journal of Pediatric Ophthalmolgy and Strabismus* 42, 2 (2005), 82–88. DOI: https://doi.org/10.3928/01913913-20050301-02

[76] Beatriz Redondo, Jesús Vera, Alba Ortega-Sánchez, Rubén Molina, and Raimundo Jiménez. 2020. Effects of a blue-blocking screen filter on accommodative accuracy and visual discomfort. *Ophthalmic and Physiological Optics* 40, 6 (2020), 790–800. DOI: https://doi.org/10.1111/opo.12738

[77] Michael Rietzler, Florian Geiselhart, Jan Gugenheimer, and Enrico Rukzio. 2018. Breaking the tracking: Enabling weight perception using perceivable tracking offsets. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, 1–12. DOI: https://doi.org/10.1145/3173574.3173702

[78] Michael Rietzler, Jan Gugenheimer, Teresa Hirzle, Martin Deubzer, Eike Langbehn, and Enrico Rukzio. 2018. Rethinking redirected walking: On the use of curvature gains beyond perceptual limitations and revisiting bending gains. In *Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality*. 115–122. DOI: https://doi.org/10.1109/ISMAR.2018.00041

[79] Mark Rosenfield. 2011. Computer vision syndrome: A review of ocular causes and potential treatments. *Ophthalmic and Physiological Optics* 31, 5 (2011), 502–515. DOI: https://doi.org/10.1111/j.1475-1313.2011.00834.x

[80] Mark Rosenfield. 2016. Computer vision syndrome (aka digital eye strain). *Optometry in Practice* 17, 1 (2016), 1–10. Retrieved from https://www.college-optometrists.org/oip-resource/computer-vision-syndrome--a-k-a--digital-eye-strain.html.

[81] Mark Rosenfield, Rui Ting Li, and Nancy T. Kirsch. 2020. A double-blind test of blue-blocking filters on symptoms of digital eye strain. *Work* 65, 2 (2020), 343–348. DOI: https://doi.org/10.3233/WOR-203086. Publisher: IOS Press.

[82] Robert Rosenthal and Donald B. Rubin. 2003. requivalent: A simple effect size indicator. *Psychological Methods* 8, 4 (2003), 492–496. DOI: https://doi.org/10.1037/1082-989X.8.4.492. Place: US Publisher: American Psychological Association.

[83] Thomas W. Schubert. 2003. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. für Medienpsychologie* 15, 2 (2003), 69–71.

[84] Sarah Sharples, Sue Cobb, Amanda Moody, and John R. Wilson. 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29, 2 (2008), 58–69. DOI: https://doi.org/10.1016/j.displa.2007.09.005. Health and Safety Aspects of Visual Displays.

[85] James E. Sheedy. 2007. The physiology of eyestrain. *Journal of Modern Optics* 54, 9 (2007), 1333–1341. DOI: https://doi.org/10.1080/09500340600855460

[86] James E. Sheedy, John Nick Hayes, and Jon Engle. 2003. Is all asthenopia the same? *Optometry and Vision Science: Official Publication of the American Academy of Optometry* 80, 11 (2003), 732–739. https://doi.org/10.1097/00006324-200311000-00008

[87] Amy L. Sheppard and James S. Wolffsohn. 2018. Digital eye strain: Prevalence, measurement and amelioration. *BMJ Open Ophthalmology* 3, 1 (2018). DOI: https://doi.org/10.1136/bmjophth-2018-000146

[88] Kien T. P. Thai, Sungchul Jung, and Robert W. Lindeman. 2020. On the use of "Active Breaks" to perform Eye Exercises for more Comfortable VR Experiences. In *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. 468–476. DOI: https://doi.org/10.1109/VRW50115.2020.00099

[89] Kazuo Tsubota. 1998. Tear dynamics and dry eye. *Progress in Retinal and Eye Research* 17, 4 (1998), 565–596. DOI: https://doi.org/10.1016/S1350-9462(98)00004-4

[90] Kazuo Tsubota and Katsu Nakamori. 1993. Dry eyes and video display terminals. *New England Journal of Medicine* 328, 8 (1993), 584–584. DOI: https://doi.org/10.1056/NEJM199302253280817. PMID: 8426634.

[91] Yury Turkov, Elena Anikina, Damien Loveland, and Faouzi Zerrouk. 2005. System and Method for Relieving Eye Strain. US Patent App. 10/488,761.

[92] Pascal W. M. Van Gerven, Fred Paas, Jeroen J. G. Van Merriënboer, and Henk G. Schmidt. 2004. Memory load and the cognitive pupillary response in aging. *Psychophysiology* 41, 2 (2004), 167–174. DOI: https://doi.org/10.1111/j.1469-8986.2003.00148.x

[93] Mindaugas Vasiljevas, Thomas Gedminas, A. Ševčenko, M. Jančiukas, Thomas Blažauskas, and Robertas Damaševičius. 2016. Modelling eye fatigue in gaze spelling task. In *Proceedings of the 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing*. IEEE, New York, NY, 95–102. DOI: https://doi.org/10.1109/ICCP.2016.7737129

[94] Cyril Vienne, Laurent Sorin, Laurent Blondé, Quan Huynh-Thu, and Pascal Mamassian. 2014. Effect of the accommodation-vergence conflict on vergence eye movements. *Vision Research* 100, 1 (2014), 124–133. DOI: https://doi.org/10.1016/j.visres.2014.04.017

[95] Roger Wagner. 1999. Anti-eye Strain Apparatus and Method. US Patent 5,933,130.

[96] George Wald. 1964. The receptors of human color vision. *Science* 145, 3636 (1964), 1007–1016. Retrieved from http://www.jstor.org/stable/1714070.

[97] Yan Wang, Guangthao Zhai, Shaoqian Zhou, Sichao Chen, Xiongkuo Min, Zhongpai Gao, and Meng-Han Hu. 2018. Eye fatigue assessment using unobtrusive eye tracker. *IEEE Access* 6, 1 (2018), 55948–55962. DOI: https://doi.org/10.1109/ACCESS.2018.2869624

[98] Barbara Wilhelm, Henner Giedke, Holger Lüdtke, Evelyn Bittner, Anna Hofmann, and Helmut Wilhelm. 2001. Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. *Journal of Sleep Research* 10, 1 (2001), 1–7. DOI: https://doi.org/10.1046/j.1365-2869.2001.00239.x

[99] Matthias Wille, Britta Grauel, and Lars Adolph. 2013. Strain caused by head mounted displays. In *Proceedings of the Human Factors and Ergonomics Society Europe*. 267–277. Retrieved from http://www.hfes-europe.org/wp-content/uploads/2014/06/Wille.pdf.

[100] J. Margaret Woodhouse and Fergus W. Campbell. 1975. The role of the pupil light reflex in aiding adaptation to the dark. *Vision Research* 15, 6 (1975), 649–653. DOI: https://doi.org/10.1016/0042-6989(75)90279-5

[101] Peggy Wright, Diane Mosser-Wooley, and Bruce Wooley. 1997. Techniques &Amp; tools for using color in computer interface design. *XRDS* 3, 3 (April 1997), 3–6. DOI: https://doi.org/10.1145/270974.270976

[102] Xingyao Yu, Dongdong Weng, Jie Guo, Haiyan Jiang, and Yihua Bao. 2018. Effect of using HMDs for one hour on preteens visual fatigue. In *Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct*. IEEE, New York, NY, 93–96. DOI: https://doi.org/10.1109/ISMAR-Adjunct.2018.00042

[103] Fabrizio Zeri and Stefano Livi. 2015. Visual discomfort while watching stereoscopic three-dimensional movies at the cinema. *Ophthalmic and Physiological Optics* 35, 3 (2015), 271–282. DOI: https://doi.org/10.1111/opo.12194

Received May 2021; accepted October 2021