National Defence Défense nationale

DEFENCE RESEARCH AND DEVELOPMENT CANADA (DRDC)
RECHERCHE ET DÉVELOPPEMENT POUR LA DÉFENSE CANADA (RDDC)

Comparing cybersickness in two levels of virtuality in a mixed reality head-mounted display

Ramy Kirollos
Wasim Merchant
DRDC – Toronto Research Centre

Tania Randall
DRDC – Atlantic Research Centre

Terms of Release: This document is approved for public release.

Defence Research and Development Canada
Scientific Report
DRDC-RDDC-2022-R158
October 2022

Canada

IMPORTANT INFORMATIVE STATEMENTS

The data collected as part of this study was approved either by Defence Research and Development Canada's Human Research Ethics Board or by the Director General Military Personnel Research & Analysis' Social Science Research Review Board.

Endorsement statement: This publication has been peer-reviewed and published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada. Inquiries can be sent to: Publications.DRDC-RDDC@drdc-rddc.gc.ca.

Abstract

A new underwater battlespace visualization capability is being developed for the Royal Canadian Navy. Mixed reality head-mounted displays may be promising technologies that are well-suited to provide 3-dimensional (3D) visualization of the underwater battlespace. However, cybersickness in *virtual* reality head-mounted displays is known to be a considerable health and safety risk and it remains unknown if cybersickness poses a similar challenge in *mixed* reality head-mounted displays. We present results from an experiment that assessed cybersickness in a mixed reality head-mounted display. We manipulated the amount of graphics or "virtuality" in two conditions to determine its effects on cybersickness: a "Mixed Reality" condition with low-level virtuality and a "Mixed Reality +" condition with high-level virtuality. Participants reported significantly greater cybersickness scores in the Mixed Reality + condition in later blocks than in the Mixed Reality condition. Moreover, participants reported significantly greater cybersickness scores as the experiment went on in the Mixed Reality + condition. We analyzed the impact of participant inter-pupillary distance, stereoacuity and history of motion sickness on cybersickness scores but results were not significant. Our findings led us to conclude that addition of graphic objects in mixed reality head-mounted displays will likely increase cybersickness, so their addition should be carefully considered. We also concluded that using mixed reality head-mounted displays with low-level virtuality produces negligible cybersickness for shore-based naval applications. Further investigation on board Royal Canadian Navy vessels is needed to assess the impact of ship motion on cybersickness levels.

Significance to defence and security

Mixed reality head mounted displays overlay and integrate graphic elements onto the real world. These displays are being considered to provide 3-dimensional visualizations of the underwater battlespace for the Royal Canadian Navy. However, it is unknown if mixed reality head-mounted displays produce nausea and other debilitating symptoms associated with cybersickness. This lab-based experiment investigated cybersickness in different conditions and found that cybersickness was minimal when limited graphic elements were presented. Therefore, mixed reality head mounted displays should be further tested for safety considerations aboard Royal Canadian Navy vessels as long as graphically presented objects are limited.

Résumé

Une nouvelle capacité de visualisation de l’espace de combat sous-marin est en cours de développement pour la Marine royale canadienne. Les visiocasques de réalité mixte pourraient être des technologies prometteuses qui conviendraient bien à la visualisation tridimensionnelle (3D) de l’espace de combat sous-marin. Cependant, l’on sait que le cybermalaise en *réalité virtuelle* est un risque considérable pour la santé et la sécurité, et l’on ignore toujours s’il pose un problème semblable dans les visiocasques en *réalité mixte*. Nous présentons les résultats d’une expérience qui a évalué le cybermalaise dans un visiocasque en réalité mixte. Nous avons manipulé le nombre de graphiques ou la « virtualité » dans deux conditions, soit en « réalité mixte » avec une virtualité faible et en « réalité mixte + » avec une virtualité élevée, afin de déterminer les effets des graphiques sur le cybermalaise. Les participants ont rapporté des scores de cybermalaise considérablement plus élevés en condition de réalité mixte + qu’en réalité mixte dans les blocs ultérieurs. De plus, les participants ont rapporté des scores de cybermalaise considérablement plus élevés au fur et à mesure de l’expérience en réalité mixte +. Nous avons analysé l’impact de l’écart interpupillaire des participants, de la stéréoacuité et des antécédents de mal des transports sur les scores de cybermalaise, mais les résultats n’étaient pas significatifs. Nos résultats nous ont amenés à conclure que l’ajout d’objets graphiques dans les visiocasques de réalité mixte augmentera probablement le cybermalaise, de sorte que leur ajout devrait être soigneusement examiné. Nous avons également conclu que l’utilisation de visiocasques de réalité mixte avec une faible virtualité provoque un cybermalaise négligeable pour les applications navales basées à terre. D’autres recherches à bord des navires de la Marine royale canadienne sont nécessaires pour évaluer l’incidence du mouvement du navire sur les degrés de cybermalaise.

Importance pour la défense et la sécurité

Les visiocasques de réalité superposent et intègrent des éléments graphiques au monde réel. Ces visiocasques sont envisagés pour fournir des visualisations tridimensionnelles de l’espace de combat sous-marin pour la Marine royale canadienne. Cependant, on ne sait pas si les visiocasques de réalité mixte provoquent des nausées et d’autres symptômes débilitants associés au cybermalaise. La présente expérience en laboratoire a permis d’étudier le cybermalaise dans diverses conditions et de constater que le cybermalaise est minime lorsque des éléments graphiques limités sont présentés. Par conséquent, les visiocasques de réalité mixte devraient faire l’objet d’autres essais pour des raisons de sécurité à bord des navires de la Marine royale canadienne, à condition que les objets présentés graphiquement soient limités.

Table of contents

List of figures

List of tables

Acknowledgements

The authors would like to thank Chris Apostoli for developing the experimental conditions, Elaine Maceda and Captain Igor Da Silva for their assistance with participant recruitment, and the participants from 32 Brigade and Canadian Forces Environmental Medicine Establishment for their time in completing this experiment.

1 Introduction

As part of Defence Research and Development's (DRDC) underwater battlespace awareness (UBA) project, DRDC – Atlantic Research Centre in partnership with Toronto Research Centre is developing a research prototype to evaluate concepts and requirements for a future underwater battle awareness tool for the Royal Canadian Navy (RCN). The eventual UBA tool will be a visualization system that provides situational awareness of the underwater battle space in support of operational decision-making by underwater warfare experts aboard RCN vessels. DRDC is assessing the possibility of a multi-display approach to the UBA prototype development. The multi-display concept being evaluated for UBA consists of a horizontal tabletop display for collaborative work, desktop displays for individual and text-heavy tasks, and mixed reality head-mounted displays (MR HMD) for 3D visualizations of the underwater battlespace. When evaluating new technologies such as the MR HMDs for the RCN, both the capabilities that these technologies offer and the possible negative impacts of using them must be considered.

An important example of negative impacts of technologies, particularly in MR HMDs, is cybersickness (CS). CS is still common and often severe in virtual reality (VR) HMDs (Palmisano et al., 2020; Rebenitsch & Owen, 2021; Yildirim, 2019). In contrast, some studies have shown that MR HMDs produce milder CS than VR HMDs (Allcoat et al., 2021; Champney et al., 2015; Moro et al., 2017; Van Benthem et al., 2021; Vrellis et al., 2020). However, many of the studies investigating MR HMDs did not directly investigate 1) the impact of MR HMDs on CS severity, and 2) the causes of CS in MR HMDs. This Scientific Report (SR) describes a DRDC – Toronto Research Centre-led investigation of MR HMD-invoked CS, which was conducted as part of the UBA team's overall evaluation of the suitability of MR HMDs for use by the RCN. In the present study, the amount of graphically rendered elements within a scene displayed in an MR HMD were manipulated across two experimental conditions while CS was indexed.

1.1 Cybersickness in extended reality head-mounted displays

1.1.1 Definitions

Augmented reality (AR) is defined as the display of virtual graphics overlaid onto the physical world. In VR, virtual graphics completely occlude the physical world. MR integrates virtual graphics with physical properties of the physical world, allowing interactive and integrated use of graphics. Extended reality (xR) is an umbrella term encompassing VR, AR, and MR technologies (He et al., 2019; Kirollos & Harriott, 2021; Merchant & Kirollos, 2022; Milgram & Kishino, 1994; XR Collaboration, 2021).

Motion sickness (MS) is malaise characterized by vomiting, retching, pallor, sweating, salivating, nausea, ocular fatigue, incapacitation, discomfort, irritability and trouble communicating (Kennedy et al., 2010). MS can arise in a variety of environments, and thus symptom and sign severity may vary slightly accordingly (Casali & Frank, 1986; Mittelstaedt et al., 2018; Rebenitsch & Owen, 2016). Airsickness, carsickness, seasickness and simulator sickness are some examples of the various types of MS resulting from different environments. MS and its various types are of particular relevance in military contexts as their associated signs and symptoms can hinder warfighter performance and compromise safety (Kirollos & Jarmasz, 2021; Proietti et al., 2021). There is debate on the definition of CS, and if it should be limited to sickness associated with head-fixed displays such as xR HMDs (Arcioni et al., 2018; Kirollos & Jarmasz, 2021), or if it should also include sickness from world-fixed displays such as 2D monitors (Gallagher & Ferrè, 2018; Kennedy et al., 2003). In this SR, we defined CS as a variation of MS that emerges from the use of xR HMDs (Arcioni et al., 2018; Kirollos & Jarmasz, 2021).

1.1.2 Brief overview of motion sickness theories

There are various theories attempting to explain, predict and reduce causes of MS. The Poison theory posits that signs and symptoms of MS are an evolutionary by-product of a toxin detection and expulsion mechanism

(Treisman, 1977). The Postural instability theory suggests that MS can be predicted based on the difficulty in maintaining postural stability by an individual before MS occurs (Riccio & Stoffregen, 1991). The neural mismatch theory explains that MS occurs because of a sustained conflict between visual and vestibular inputs (Reason, 1978) and is the most accepted theory of MS. Rest frame hypothesis (RFH) predicts that in the absence of a stable cue to the horizon (i.e., the point at which the sky and the earth's surface appear to meet), MS becomes proportionally severe (Prothero & Parker, 2003). RFH is arguably related to neural mismatch theory as it infers that lack of cues for spatial orientation aggravates the visual-vestibular conflict (Hemmerich et al., 2020). However, when a clear and reliable reference cue such as the horizon is present, sickness can be far less severe as these provide relative spatial orientation cues to an observer. Despite the many attempts at explaining and mitigating MS and CS, there is still no universally accepted agreement or solution that explains and that can fully mitigate MS (Lawson, 2014).

1.1.3 Overview of measures of cybersickness

There are many self-report, behavioural and physiological measures to index sickness susceptibility and sickness state (see Merchant and Kirollos (2022) for a recent review). Motion sickness susceptibility methods are used to index past episodes of sickness. Examples of these include the short motion sickness susceptibility questionnaire (MSSQ) and the long MSSQ (Golding, 1998, 2006). Some research has shown that past sickness can predict present sickness, providing some evidence that some sickness susceptibility questionnaires significantly correlate with sickness state questionnaires (Beadle et al., 2021; Golding et al., 2021).

Sickness state questionnaires can be used to determine how sick an individual is in real time. Examples of sickness state questionnaires include the simulator sickness questionnaire (SSQ), the fast motion scale (FMS) and the virtual reality sickness questionnaire (VRSQ) (Kennedy, Lane, et al., 1993; Keshavarz & Hecht, 2011; Kim et al., 2018). Among all of these, the SSQ is the most well-established and well-validated sickness state questionnaire (Kemeny et al., 2020; Weech et al., 2019) and was used in this study. The SSQ provides a total severity score (TS) of approximately 235 (Bimberg et al., 2020). The TS is made up of three subscales. The questions on the SSQ that make up the subscales are not mutually exclusive. These subscales are: nausea (N), ocular discomfort (O) and disorientation (D) scores.

Researchers have tried to develop reliable physiological measures of sickness (Gavgani et al., 2018; Hemmerich et al., 2020; Jang et al., 2022; Weech et al., 2020). These methods rely on neurophysiological correlates of MS, heart rate, cortisol levels, and galvanic skin conductors. However, to date, these methods are still not precise enough to reliably detect the presence of sickness across individuals. Postural sway is a behavioural measure used to attempt to predict sickness but there is debate in the literature as to whether it is a reliable sole indicator of sickness (Arcioni et al., 2018; Gower et al., 1988; Reed-Jones et al., 2008; Stoffregen & Smart Jr, 1998). Thus, self-report measures, and the SSQ in particular, continue to be a highly relied upon method to assess sickness.

1.1.4 Contributors to cybersickness

There are many known factors contributing to CS when using MR HMDs. Some of the factors that contribute to CS include visual conflicts, visual-vestibular conflicts and individual differences.

1.1.4.1 Visual conflicts

A visual conflict that contributes to symptoms of CS involves the proximity of the display to the eyes in combination with artificial manipulation of depth cues in graphically rendering elements called the vergence accommodation mismatch (VAM) (Hoffman et al., 2008). Another visual conflict that contributes to CS is a mismatch between an individual's interpupillary distance (IPD) and the IPD set in the HMD. Stanney et al. (2020) found that HMD IPD setting that differs from an individual's IPD contributes to CS.

1.1.4.2 Visual-vestibular conflicts

There are three known visual-vestibular conflicts contributing to CS severity: motion to photon lag, vection, and lack of rest frames.

1.1.4.2.1 Motion to photon lag

Motion to photon lag is characterized by the lag between the image displayed according to a previous head position to the image displayed corresponding to an updated head position. This directly contributes to CS because there is a delay between what is expected to be seen and what is actually seen based on the new head position. For the most part, native lag of an xR HMD for yaw axis head movement, the most common type of head movement, is increasingly negligible (Kim et al., 2020). Kim et al. (2020) manipulated VR HMD lag during yaw axis head movement. They found that CS scores were negligible (i.e., below a score of 1/20 on the FMS) in their native lag condition (~4 ms).

1.1.4.2.2 Vection

Another visual-vestibular conflict contributing to CS is implied self-motion while stationary, commonly known as vection (e.g., the sense that a *stationary* train an observer is on is moving backwards as an adjacent train moves forward). Although there is debate as to whether vection is a prerequisite for CS and other types of MS, vection can generate visual-vestibular conflict because the visual system perceives self-motion while the vestibular system does not (Kuiper et al., 2019; Lawson, 2005; Risi & Palmisano, 2019; Teixeira et al., 2021).

1.1.4.2.3 Lack of rest frames

Lack of rest frames (RFs) is another visual-vestibular conflict that contributes to CS. In MR HMDs, the graphic content is overlaid on fixed elements of the physical environment. These visual properties of the physical world should serve as RFs as they act as stable orientation cues relative to an observer. Moro et al. (2017) compared the use of VR HMDs, MR HMDs and tablets as training tools for learning anatomy. They found that participants tended to report symptoms of CS when using VR compared with MR HMDs and tablets. This finding makes sense with regards to RFH because VR occludes the physical world completely and therefore provides no real-world RFs in contrast to MR HMDs. A caveat of this study was that authors did not use a recognized measure of CS. Work by Kemeny et al. (2017), Cao et al. (2018), Hemmerich et al. (2020) and Whittinghill et al. (2015) all found that RFs reduced CS. Thus, CS is thought to be less severe in MR compared to VR because MR devices permit the viewer to perceive a physical reference cue (Kuiper et al., 2019; Lawson, 2005; Risi & Palmisano, 2019).

1.1.4.3 Other considerations

Exposure time has been shown to contribute significantly to CS (Hemmerich et al., 2020; Jasper et al., 2020; Lawson et al., 2021; Palmisano et al., 2020; Porcino et al., 2021). Individual factors including stereoacuity, ethnicity, gender, familiarity with HMDs, history of MS and others may contribute to CS in ways that are not yet clear (Stanney et al., 2020). Finally, there may still be unknown contributors to CS.

Van Benthem et al. (2021) recently performed a review of CS when using MR HMDs. They found a few studies showing that CS was milder in MR than in VR. However, they found that few studies have reported specifically on CS in MR HMDs and even fewer studies focused on CS resulting from MR HMDs. The report by Van Benthem et al. concluded that studies using MR HMDs generally indicated less severe CS than studies employing VR HMDs. However, a thorough and focused investigation of CS in MR HMDs is still required.

1.2 Present study

We have identified a gap in the literature wherein the severity of CS in MR HMDs has not been directly investigated. To fill this research gap, an experiment was conducted using MR HMDs. The amount and type of graphically generated elements in the scene were manipulated in this experiment to determine their effects on CS. Two conditions were evaluated: in the first condition only foreground objects were graphically generated. This was called the “MR” condition. In the second condition, foreground and background objects were graphically generated. This was the “MR+” condition. The MR+ condition therefore simulates VR to a large extent as the entire visual scene is graphically generated. Findings from this study will inform the interface display technologies chosen for UBA’s prototype by determining the impact of graphic quantity on CS.

1.2.1 Hypotheses, assumptions and possible outcomes for UBA

Hypothesis 1—increased virtuality in MR HMDs will lead to increased CS over time. We modulated graphic objects across the MR and MR+ condition whereby the MR+ had more graphically rendered elements in the scene than the MR condition. The MR condition had more RFs than the MR+ condition. We hypothesized that CS scores would be greater for the MR+ condition than the MR condition over time.

Hypothesis 2—increased MR HMD exposure time will lead to increased CS in each condition. Many studies have consistently shown that exposure time is a main predictor of sickness severity (Hemmerich et al., 2020; Jasper et al., 2020; Lawson et al., 2021; Palmisano et al., 2020; Porcino et al., 2021). Therefore, we hypothesized that participants would experience greater CS in the final block of the experiment compared to the first few blocks.

Hypothesis 3—greater susceptibility to motion sickness will lead to increased CS. Some researchers have found a significant correlation between history of motion sickness, measured by the motion sickness susceptibility questionnaire (MSSQ) and current sickness measured by the simulator sickness questionnaire (SSQ) (Beadle et al., 2021; Golding et al., 2021). We aimed to replicate this finding and therefore hypothesized that participants with high MSSQ scores would have high SSQ scores.

VAM was expected to contribute to CS in this experiment consistent with reports by Hoffman et al. (2008) on the contribution of VAM to increased CS symptoms. It was predicted that there would be greater CS scores for participants in the MR+ condition than in the MR condition as a result of greater VAM in the MR+ condition caused by the greater amount of graphically rendered elements in the MR+ condition. However, we did not directly isolate or control for VAM and as such, this was not a hypothesis.

IPD difference between the HMD and participant should not have contributed to sickness in this experiment because HMD IPD was calibrated to each individual’s IPD. The display presented a stationary environment, ruling out the possibility of CS occurring from visual-vestibular conflict that could arise from vection. Participants moved their heads in the yaw axis, also limiting the contribution of visual-vestibular conflict (specifically, motion to photon lag) to CS as this is known to have a negligible effect (Kim et al., 2020). We therefore assumed that HMD IPD, head movement, and vection would have only negligible contributions to CS in this experiment.

We outlined three possible outcomes regarding what results from this lab-based experiment could mean with regards to our recommendations for at-sea testing and UBA: 1) If CS scores in the MR condition were low in later blocks, but high in the MR+ condition, this would support recommendation of MR HMDs with low virtuality for further investigation aboard RCN vessels. 2) If CS scores are low in the MR and MR+ conditions, this will support recommendation to use MR HMDs with either low or high virtuality for further investigation aboard RCN vessels. 3) If CS scores are high in the MR and MR+ conditions, this will discourage recommendation of MR HMDs for further investigation aboard RCN vessels and for UBA prototyping. We did not anticipate that CS would be higher in the MR condition than in the MR+ condition.

2 Method and materials

2.1 Participants

Participants were healthy Canadian Armed Forces (CAF) members with normal or corrected-to-normal vision assessed with a screening question in the demographic questionnaire and no history of neurological or vestibular disorder (n = 24, M_{age} = 32, SD_{age} = 9.9). Participants could wear their spectacles while wearing the MR HMD. All participants were recruited from 32 Brigade and Canadian Forces Environmental Medicine Establishment. Participants were briefed on the experiment, provided informed consent and were free to withdraw from the study at any point during experimentation. Exclusion criteria for the study included any uncorrected visual deficits, balance/vestibular disorders, and neurological conditions. Remuneration was provided to all participants. Three participants in the experiment were naval personnel. Two participants had limited experience aboard naval vessels while the other had extensive experience onboard navy vessels. Three participants were female.

2.2 Materials

An in-house application was developed in Unity game engine (version 2019.2), a cross-platform game engine created by Unity Technologies. The application was designed to generate visual stimuli for conditions of the experiment through custom scripts written in C#.

The Microsoft HoloLens 2 MR HMD was used to display visual stimuli in all experimental conditions. The HoloLens 2 can render graphics overlaid onto the physical environment and allows for viewing of virtual objects in 3D. This HMD has a native 1440 x 936 pixels resolution per eye, 60 Hz refresh rate, and 43° horizontal by 29 vertical field of view (FOV) per eye.

2.3 Stimuli and conditions

The HoloLens 2 displayed the custom application with one of two experimental conditions. These scenes are 3D renditions of an office, including graphical objects such as a desk, chair, computer, etc. The dimensions of the virtual office were 2.72 m wide, 3.62 m in length and 2.57 m from the floor to the ceiling. The “MR” condition displayed graphically generated foreground objects such as furniture and portraits hung on the virtual room’s wall, while background objects such as the walls and ceiling were not represented graphically. In the “MR+” condition, all objects in the MR condition as well as the walls were graphically rendered. Participants were randomly assigned to either the MR condition, or the MR+ condition represented in Figure 1.

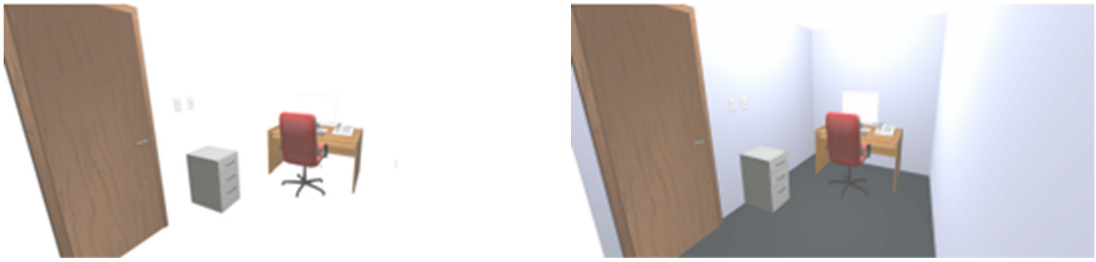

Figure 1: *Simplified approximations of the MR condition (left) and the MR+ condition (right).*

2.4 Design

This experiment had a mixed factors design. The between-subjects factor was visual condition with two levels: MR and MR+. The within-subjects factor was *time*, with six levels: Time 1–6. Each level represents an increase in cumulative exposure time in 5-min increments. The primary dependent variable was CS, indexed with the SSQ (Kennedy, Berbaum, et al., 1993).

2.5 Procedure

Participants were tested individually during a single session. Participants provided informed consent, had their IPD and stereoacuity measured by the experimenter, completed a demographic questionnaire, MSSQ and baseline SSQ. Participants were instructed to sit upright comfortably while the experimenter demonstrated the task and introduced the participant to Microsoft HoloLens 2 MR HMD. The participant was then instructed to don the HMD while the experimenter confirmed the correct fit and placement of the device on the participant's head. Next, the experimenter used the Microsoft HoloLens Companion Application on a laptop to initiate calibration for IPD ensuring that the IPD in the HMD matched that of the participant. After successful IPD calibration, the participant was instructed to place their head in a chinrest at a designated height to render elements of the virtual scene at the appropriate height relative to the ground. The experimenter then started the custom application which used the participant's head position to determine where to display elements relative to the room's walls, floor and ceiling. The room lights were adjusted to be dim at approximately 2.5 Lux using a light meter (Extech Instruments Light Meter, LT300). A 30-second practice session took place to familiarize the participant with the experimental task. Once the participant demonstrated that they understood the task and had no further questions, the first of five, 5-min blocks of the experiment began. In both the MR and MR+ conditions, participants viewed a virtual representation of an office room and made yaw axis head movements to the left and right while seated. They were instructed to look at the virtual content that was in line with their head orientation as they moved. The sound of a metronome generated by the HMD every 3 s cued participants to make angular yaw axis movements with their heads from left shoulder to right shoulder repeatedly throughout the block demonstrated in Figure 2.

Figure 2: *A participant performing the experimental task of turning their head left-to-right at the 3-s constant interval of an auditory cue while wearing the MR HMD.*

The SSQ was administered via paper and pen prior to the participant donning the HMD and at the end of each block during the 1-min break. Overall, six SSQ scores were obtained from each participant. Following the response to the final SSQ, the HMD was removed.

2.6 Measures

2.6.1 Visual assessment tests

The Random Dot Stereo Acuity Test (Stereo Optical Company Inc., Stereo Acuity Test Version 2012) was used to assess stereoacuity. IPD was measured using the Reichert PDM Digital PD Meter prior to experimentation (Model No. 15020).

2.6.2 Demographic questionnaire

The demographic questionnaire recorded each participant sex, age, experience on Navy vessels, video game and simulator experience, recent use of any substance that can impact nausea (e.g., anti-nauseogenics medications, alcohol, recreational drugs), history of neurological and vestibular disorders as well as any visual impairments. The demographic questionnaire is attached in Annex A.

2.6.3 Motion sickness susceptibility questionnaire

The short form of the MSSQ (Golding, 2006) was used to measure participant susceptibility to motion sickness (MS). The MSSQ was administered prior to experimentation to capture participant's previous experiences with MS when using different modes of transportation, both as a child (MSSQ A) and as an adult (MSSQ B). Participants rated each transportation item on a scale of 0 (*never got sick*) to 3 (*often got motion sick*) on the MSSQ.

2.6.4 Simulator sickness questionnaire

The SSQ (Kennedy, Lane, et al., 1993) contains 16 symptoms (e.g., nausea, fatigue, discomfort). The intensity of each symptom is rated by participants on a four-point Likert scale *(not at all = 0*, *mild = 1*, *moderate = 2,* and *severe = 3*). The SSQ is divided into 3 subscales to measure the three most common symptoms of motion sickness: Nausea (N), Ocular discomfort (O) and Disorientation (D). The TS is computed as the weighted average of the sum of these three subscales multiplied by a constant value (see Kennedy et al., 1993; and Merchant and Kirollos, 2022 for further details on the SSQ).

According to Kennedy et al. (2003), a TS greater than 20 on the SSQ is indicative of a "problem simulator" and should be avoided. The authors further state that TS scores of 15–20 suggest that symptoms are concerning, and SSQ TS scores of 10–15 are significant symptoms and any score below that to be of little or negligible concern.

3 Results

SSQ data comparing MR to MR+ results were analyzed using the Kruskal-Wallis test because this is a non-parametric statistical test that can be used for between-subjects data. Within-subjects SSQ data comparing MR and MR+ data across blocks were analyzed using Friedman's test because this is a non-parametric statistical test that can be used for within-subjects data. A Spearman correlation comparing SSQ scores to MSSQ scores was performed to determine if there was any relationship between history of sickness and sickness state. A Pearson correlation was carried out to investigate differences between IPD and stereoacuity. A Kruskal-Wallis test was used to compare SSQ scores according to stereoacuity categorization. These analyses are described below.

3.1 SSQ TS: Scores

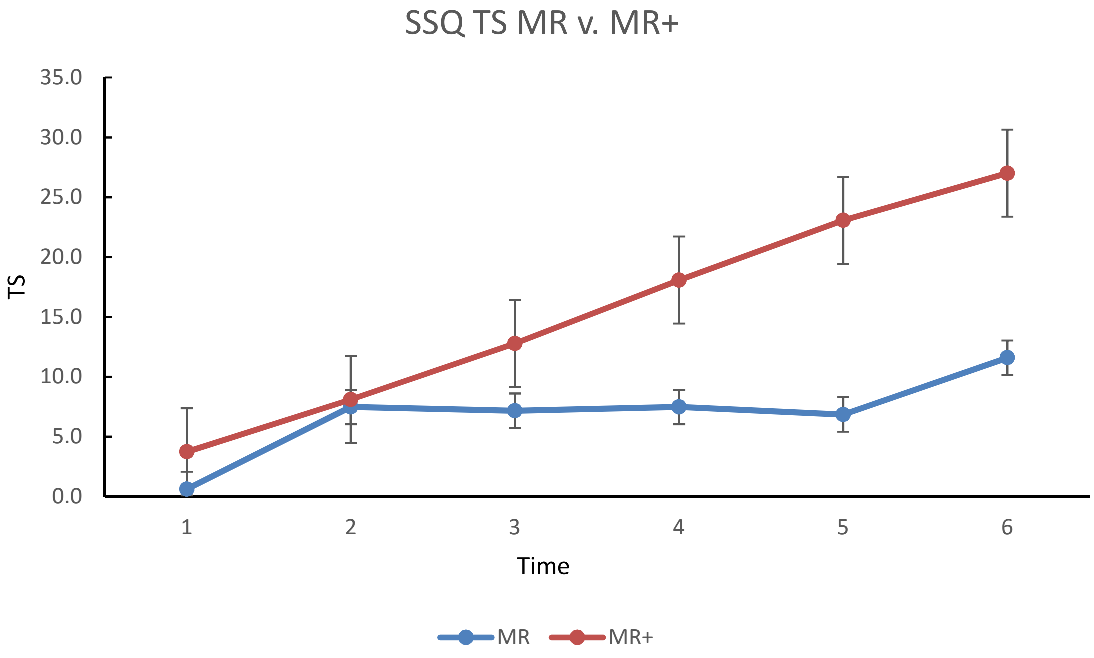

Figure 3: *Mean TS scores across participants from baseline (1) to Time 6. TS scores from the MR condition are in blue. TS scores from the MR+ condition are in red. Error bars represent standard error rates.*

The two primary research questions we aimed to address with our analyses were to determine: 1) differences in SSQ scores between MR and MR+ conditions in later blocks (Hypothesis 1), and 2) impact of exposure time on SSQ scores (Hypothesis 2). Figure 3 above represents these data.

Five Kruskal-Wallis tests were performed to compare SSQs scores from Times No. 2–6 for the MR condition with the corresponding SSQs for the MR+ condition (e.g., MR Time No. 3 v. MR+ Time No. 3) to determine if there was a significant difference between SSQ TS scores of each condition. Effect size for Kruskal-Wallis test was calculated by determining Epsilon squared ($\mathcal{E}^2$) (Tomczak & Tomczak, 2014). Findings indicated a significant difference between TS scores of the MR v. MR+ condition in Time No. 5 (after 20 minutes), $H(1) = 6.888$, $p = 0.009$, $\mathcal{E}^2 = 0.299$. This demonstrates a relatively strong effect size according to Rea and Parker (2014). A significant difference between TS scores of the MR v. MR+ condition in Time No. 6 (after 25 minutes) was also found, $H(1) = 4.255$, $p = 0.039$, $\mathcal{E}^2 = .185$. This demonstrates a relatively strong effect size according to Rea and Parker. All other comparisons were not significant and are displayed in Table 1. These findings indicate that there is a significant difference in TS scores between MR and MR+ conditions for the later Time blocks.

Table 1: *Statistical output of Kruskal-Wallis tests for all comparisons.*

Comparison	Statistical output
MR Time 2 v. MR+ Time 2 (After 5 minutes of exposure)	H(1) = .926, $p = 0.336$
MR Time 3 v. MR+ Time 3 (After 10 minutes of exposure)	H(1) = 2.787, $p = 0.095$
MR Time 4 v. MR+ Time 4 (After 15 minutes of exposure)	H(1) = 3.183, $p = 0.074$
MR Time 5 v. MR+ Time 5 (After 20 minutes of exposure)	H(1) = 6.888. $p = 0.009$
MR Time 6 v. MR+ Time 6 (After 25 minutes of exposure)	H(1) = 4.255, $p = 0.039$

To analyze the difference across SSQ TS scores across Times 1–6, our within-subjects effect, a Friedman test was performed on the MR data and a second Friedman test was performed on the MR+ data. The Friedman test comparing the 6 Time points in the MR condition indicated no significant differences, $X^2 (12) = 10.28, p = 0.068$. The Friedman test comparing the 6 Time points in the MR+ condition indicated a significant difference. Effect size was calculated with Kendall's W (W), $X^2 (12) = 31.176, p <.001$, W = 0.520. According to Rea and Parker (2014), this represents a strong effect size. A pairwise Wilcoxon signed rank test with a Bonferonni correction was used to compare blocks. Nine post-hoc comparisons of interest were made to compare the scores from the 6 Time points in the MR+ condition. Therefore, the Bonferonni correction employed a p cut-off value of $p = .05/9 = 0.0055$. Results indicated that there were significant differences between SSQ 1 v. SSQ 6 ($p = 0.004$), SSQ 3 v. SSQ 5 ($p = 0.004$), and SSQ 3 v. 6 ($p = 0.004$). Overall, these findings indicate that there is a significant difference and strong effect size in TS scores over Time in the MR+ condition. Moreover, Mean TS scores in the MR+ condition at Times 5 and 6 (after 20 and 25 minutes) exceed 20, indicating that in the MR+ condition, MR HMDs are categorized as problem simulators according to Kennedy et al. (2003).

3.2 SSQ subscales: N, O, D Scores

Across the MR and MR+ conditions, we separated SSQ scores into the N, O and D subscales. Data for these are presented in Figure 4. Visual assessment of data presented in Figure 4 indicate that O and D symptoms were scored higher on the SSQ than the N symptoms respectively in the MR and MR+ conditions. To determine if this difference between O and D subscales v. N were significant, we performed three Friedman tests comparing N, O and D subscales in the MR and MR+ condition collapsed over Time. We were also interested in determining if subscale scores varied as a result of condition and therefore performed two Kruskal-Wallis tests comparing each subscale score across MR and MR+ conditions (e.g., $O_{(MR)}$ v. $O_{(MR+);}$ $D_{(MR)}$ v. $D_{(MR+)}$) collapsed over Time. None of the comparisons were significant and are displayed in Table 2. These data indicate that there were no significant differences across SSQ subscales within the MR and MR+ conditions.

Table 2: *Various analyses on SSQ N-O-D subscales and their statistical outputs in analyses collapsed over Time.*

Comparison (Statistical test performed)	**Statistical output**
N-O-D Comparison in MR condition (Friedman Test)	X(12) = 0.839, $p = 0.657$
N-O-D Comparison in MR+ condition (Friedman Test)	X(12) = 2.087, $p = 0.352$
N subscale in MR v. MR+ conditions (Kruskal-Wallis Test)	H(12) = 3.087, $p = 0.079$
O subscale in MR v. MR+ conditions (Kruskal-Wallis Test)	H(12) = 3.804, $p = 0.079$
D subscale in MR v. MR+ conditions (Kruskal-Wallis Test)	H(12) = 0.296, $p = 0586$

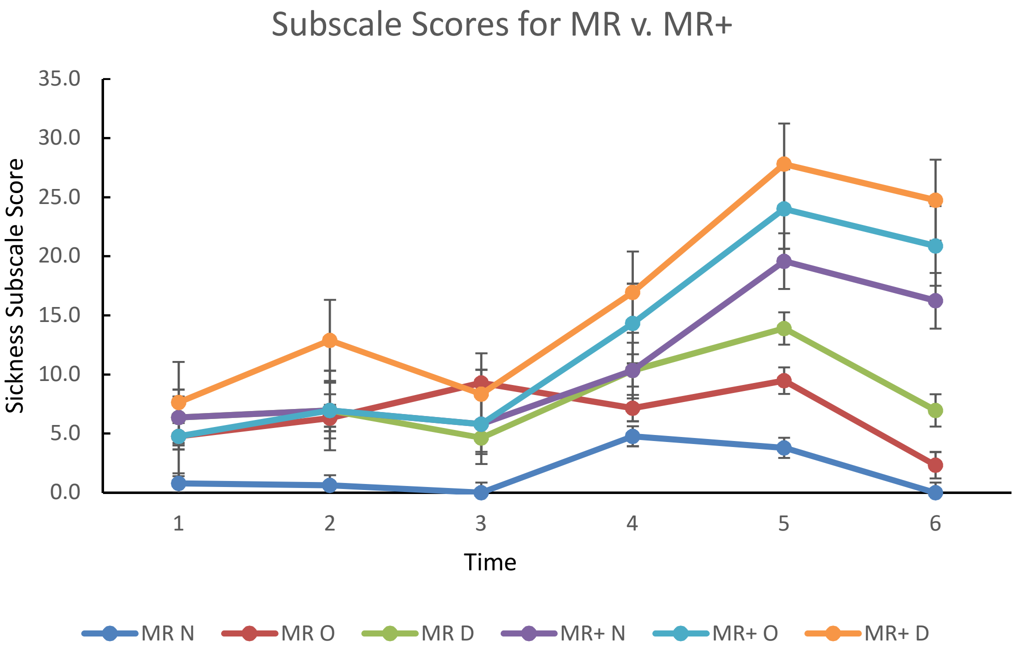

Figure 4: *Mean scores for N, O, and D subscales of the SSQ across Time for the MR and MR+ conditions. Error bars represent standard error rates.*

In a final comparison, we were interested in determining which subscales may be driving the difference between SSQ TS scores in Times 5 and 6. We therefore analyzed subscale scores within times (e.g., MRtime5(O) v. MR+time5(O)). There was a significant difference in the N subscale between MR and MR+ in Time 5, (H(12) = 4.45, $p = 0.035$, $Ɛ^2 = 0.193$) indicating a relatively strong effect size according to Rea and Parker (2014) and in the O subscale in Time 5 (after 20 minutes), H(12) = 4.94, $p = 0.026$, $Ɛ^2 = 0.215$, also indicating a relatively strong effect size. All other comparisons were not significant ($p > 0.05$).

3.3 Correlational analysis between MSSQ and SSQ

We examined the relationship between previous sickness (total MSSQ) and sickness state (mean SSQ TS) in the MR and MR+ conditions separately with two Spearman correlations. For both MR and MR+ conditions, there were no significant correlations between total MSSQ scores and SSQ TS scores (MR: $\rho(12) = 0.048$, $p = 0.882$; MR+: $\rho(12) = 0.045$, $p = 0.888$).

3.4 Impact of stereoacuity on SSQ scores

We categorized participant stereoacuity results as either “good” (<25 arc sec (s), n = 9) or “bad” (>25 s, n = 15) approximating categorization by Deepa et al. (2019) and compared both groups’ mean TS scores in a Kruskal-Wallis test to determine the impact of stereoacuity on CS. Results indicated there was no significant difference in CS between “good” and “bad” stereoacuity groups, $H(1) = 0.014$, $p = 0.905$. Further, we performed a Spearman correlation between stereoacuity and mean TS scores to determine if there was a relationship between stereoacuity and CS. Results indicated no significant correlation between stereoacuity and mean TS score, $\rho(24) = 0.121$, $p = 0.572$.

Finally, we investigated the relationship between participant IPD (M = 63.3 mm, SD = 3.7 mm) and stereoacuity (M = 163.4 s, SD = 232.8 s). A Pearson Correlation revealed that there was no significant relationship between IPD and stereoacuity, $r(24) = -0.103$, $p = 0.631$.

4 Discussion

A literature review by Van Benthem et al. (2021) indicated that research on MR HMDs did not directly examine the impact of MR HMD use on CS and the little research on this topic suggests that MR HMDs generally produce little CS. However, based on that report we identified that there has not been a thorough analysis thus far on CS with MR HMDs. Here we have directly examined the impact of MR HMDs on CS and expanded findings by Van Benthem and colleagues indicating that MR HMDs can produce CS in high virtuality HMD presentations. In this experiment, the primary objective was to determine the impact of MR HMDs on CS over time (see Hypothesis 1). We did so by varying the level of virtuality in an MR HMD and indexing CS. We selected a mixed design whereby participants only completed either the MR or MR+ condition as some studies have reported evidence of adaptation effects when participants are exposed to similar display and protocol (Beadle et al., 2021; Howarth & Hodder, 2008). In comparing results in the MR to MR+ condition, we found significant differences in SSQ scores in Times 5 and 6 (after 20 and 25 minutes) where the effect size was relatively strong. Figure 3 indicated that MR+ SSQ TS scores were higher than MR SSQ scores in all blocks. Overall, results comparing SSQ scores in the MR to MR+ condition are consistent with our hypothesis that MR+ SSQ scores would be greater than MR SSQ scores in later blocks. These results are also consistent with RF hypothesis in that the condition with fewer available rest frames (MR+) generated more severe sickness than the condition with more available rest frames (MR). These results indicate that generating graphic scenes that occupy the entire visual field in MR HMDs will produce CS. Thus, the amount of graphically rendered elements in MR HMDs must be limited in order to mitigate CS.

According to Kennedy et al. (2003), a TS greater than 20 on the SSQ is indicative of a “problem simulator” and should be avoided. The authors further state that TS scores of 15–20 suggest that symptoms are concerning, and SSQ TS scores of 10–15 are significant symptoms and anything below that to be of little or negligible concern. Our results in the MR+ condition indicate that peak mean TS scores exceeded 20 after Time 5 (after 20 minutes) (Mean $TS_{(time5)} = 23.1$, Mean $TS_{(time6)} = 27.0$) which according to Kennedy et al. (2003) qualifies as a “problem simulator” that should be avoided. In contrast, at the end of the experiment, in the MR condition, TS scores had a mean of 11.6. Therefore, based on these findings we conclude that MR HMDs are safe for use if there are only limited graphically displayed elements.

The second objective of this experiment was to determine the impact of duration wearing an MR HMD on CS (see Hypothesis 2). We performed two within-subjects analyses on SSQ scores: one for the MR condition and one for the MR+ condition. Only the MR+ condition indicated that time significantly impacted SSQ scores. This provides further evidence that saturating the MR HMD with graphics can produce severe CS over time. On the other hand, results from this comparison indicate that prolonged use of an MR headset with limited graphics should not produce severe CS. Many studies on the use of VR HMDs have indicated that CS severity increases as a function of time, consistent with our findings in the MR+ condition (Hemmerich et al., 2020; Jasper et al., 2020; Lawson et al., 2021; Palmisano et al., 2020; Porcino et al., 2021). Most of these studies report peak CS within 10–15 min of the experiment. In contrast, in our study, participants used the MR HMD for a total of 25 min and results indicated that CS increased significantly over the 25 min session of the MR+ condition. Based on these findings, we recommend use of MR HMDs with limited graphical elements for up to 25 min, and further anticipate longer durations of use to be acceptable as well.

Recently, Beadle et al. (2021) and Golding et al. (2021) found significant positive correlations between MSSQ scores and SSQ TS scores (see Hypothesis 3). These findings make sense as they indicate that participants that often felt sick previously were most likely to feel sick when using VR HMDs. Thus, we performed a correlational analysis on MSSQ data and SSQ data. However, our results did not replicate those of Beadle et al. and Golding et al. as we did not find a significant correlation between MSSQ and SSQ scores. Our findings indicate there is no evidence for a correlation between past motion sickness and present CS with MR HMDs.

Hoffman et al. (2008) created a scale to assess visual fatigue from VAM. Some of the questions on their scale deal directly with ocular fatigue and headache, thereby overlapping with questions in the SSQ. However, Hoffman et al. do not report directly on CS or any form of MS. We have not come across any research that directly investigated the relationship between any form of MS and VAM. This is likely due to the fact that VAM is a difficult conflict to isolate since it requires a headset to physically alter its focal distance, which is unpractical with commercial off the shelf HMDs that we are aware of. In our experiment, we did not isolate VAM as it was not the primary research concern though we stated that VAM would likely be greater in the MR+ condition as there were more virtually rendered elements, more depth cues and therefore more VAM. Accordingly, we found greater CS in the MR+ condition than the MR condition though it is not clear what the independent contributions of RF and VAM are to this result. It could be the case that greater VAM resulted in greater CS in our experiment. However, further research must be conducted to isolate the contribution of VAM from RF on CS severity. Based on our findings and the literature, the relationship between VAM, stereoacuity and CS in unclear and further research is needed to investigate their relationship.

To investigate if stereoacuity impacted CS, we categorized participants into two groups: “good” (less than 25 s) and “bad” (greater than 25 s) stereoacuity approximating Deepa et al. (2019) stereoacuity categorization. We compared both group’s SSQ TS scores and did not find a significant difference. Similarly, Arcioni et al. (2018) investigated stereoacuity on CS and found no relationship. Our findings and those by Arcioni et al. indicated that stereoacuity does not appear to impact CS. A recent study by Luu et al. (2021) indicated that participants with stereopsis experience more severe CS than participants without stereopsis (consistent with results in Palmisano et al. [2019]). This may initially appear to contrast our study in which “good” and “bad” stereoacuity did not impact CS. However, some important differences exist between our study and Luu et al.’s (2021). First, stereopsis which was investigated in Luu et al. (2021), is the ability to perceive in depth. This differs from stereoacuity in our study because stereoacuity is a measure of how *well* one can perceive in depth. Second, vection correlated with sickness in their study. Therefore, it is not clear if vection resulted in CS, which has been sometimes shown to increase CS, or if stereopsis solely contributed to increased CS. Third, the authors do not mention if an adjustment to the FOV in the monocular v. stereopsis condition was made or not, as larger FOVs typically produce more severe CS (Lin et al., 2002). Fourth, Luu et al. used VR HMD in contrast to our MR HMD, which creates a further challenge in comparing our results to their study. Finally, Luu et al. employed the FMS to index CS, whereas we used the SSQ in our study.

We performed various post-hoc analyses including SSQ subscale score comparisons, correlational analyses exploring the relationships between SSQ and MSSQ scores, stereoacuity and IPD, stereoacuity and SSQ scores as well as comparisons of SSQ scores of “good” stereoacuity to “bad” stereoacuity. We discuss their important implications herein.

There is ongoing debate in the literature on motion sickness, its different forms and its associated severity. For these reasons, we believed it to be important to compare SSQ subscales, N, O, D to determine if MR HMDs provide a distinct pattern of results. Upon visual inspection of Figure 4, the O and D scale had slightly greater scores than the N subscale for both the MR and MR+ conditions. However, our analysis indicated that these differences were not significant. Stanney et al. (2003) found that cybersickness from VR HMDs produces greater D subscale scores than N and O subscale scores. By contrast, they found that simulators produce greater O subscale scores than N and D subscale scores. In the present study using MR HMDs, we found that D scores followed by O scores were higher than N scores but this was not significant. Therefore, subscale scores produced by MR HMDs appear to differ from those reported in simulators and VR HMDs in Stanney et al. (2003) study though the pattern of subscale scores in MR HMDs were not significant.

We were also interested in determining if there was a correlation between IPD and stereoacuity. Some research demonstrates a relationship between IPD and stereoacuity where greater IPDs are associated with better stereoacuity (Aslankurt et al., 2013; Eom et al., 2013). This makes sense mathematically as a greater separation between the two eyes should yield stronger disparity between the left and right retinal images, and thus, greater depth perception results demonstrated by greater stereoacuity scores. However, we found one study demonstrating that smaller IPD resulted in better stereoacuity (Shafiee et al., 2014) and two studies demonstrating that IPD had no impact on stereoacuity (Arcioni et al., 2018; Mai & Schlueter, 2010). A correlation did not demonstrate a significant relationship between IPD and stereoacuity scores in our study, consistent with findings by Arcioni et al. (2018) and Mai and Schlueter (2010). However, stereoacuity scores were highly variable in our study likely contributing to our finding. These findings led us to conclude that the relationship between IPD, stereoacuity and sickness is not clear.

4.1 Recommendations

Based on results from our experiments and the above discussion, we provide the following recommendations regarding the use of MR HMDs for UBA:

1. Use of MR HMDs with limited graphically generated elements is recommended for further investigation aboard RCN vessels as these provoke no or negligible CS in the current experiment. Complete saturation of MR HMDs with graphic elements is not recommended as it can cause severe CS.

2. Provided users are employing the latest MR HMDs, users may make head movements without becoming sick as our experiment included head movements and demonstrated limited CS, provided users have affixed the HMD to their heads correctly. However, when possible, users should limit head movement to limit its minor contribution to CS.

3. MR HMDs with limited graphically generated elements are recommended for use for approximately 25 min as tested here and can safely be used for longer durations if needed.

4. IPD must be calibrated for each user in the HMD to minimize CS.

5. Users should remove the MR HMD or flip the holographic display lenses up away from view when not using it to avoid the impact of long duration use of MR HMDs on CS provided the MR HMD can be moved away from view.

4.2 Future planned research to inform UBA and CAF

In this experiment, we explored the relationship between graphically rendered elements in MR and CS. Future research will aim to mimic the MR+ condition in a VR HMD. This will allow us to compare VR under the same viewing conditions and protocol as this SR (e.g., identical FOV, scene, experimental protocol, duration, measures, etc.) to better isolate the differences in CS across VR and MR HMDs.

In this experiment, all conditions were performed on land. However, UBA will be used at sea aboard RCN vessels. These vessels will be in motion likely provoking seasickness. Therefore, we will extend our findings from this SR by performing this experiment at sea to examine if there are interactions between seasickness and CS in a future experimental campaign. Future experiments will also include RCN personnel with seasickness familiarity to investigate the relationship between CS and seasickness familiarity.

Many unplanned exploratory post-hoc analyses were performed to elucidate relationships between, IPD, stereoacuity, and CS in this SR. Most of these analyses were not significant. It is possible that sample sizes in groups were too small to reach required power thresholds for significance in these analyses. Future experiments outlined here will generate a larger data set in addition to the current data set that will allow us to repeat these post-hoc analyses and more confidently determine the impacts of IPD and stereoacuity on CS. All of these future research initiatives are planned in the next phases of our research program.

5 Conclusion

The purpose of this SR was to understand CS in MR HMDs to inform their usage for UBA. We performed an experiment modulating graphically rendered elements in MR HMDs and indexed participant CS over the course of the experiment. Our findings indicated that the high virtuality condition produced significantly greater CS than the low virtuality condition in later blocks. Moreover, CS became significantly more severe over time in the MR+ condition, qualifying it as a problem simulator to be avoided (Kennedy et al., 2003). We did not find any significant differences in the post-hoc analyses looking at individual differences such as IPD or stereoacuity. Our results indicate that graphically rendered objects in MR HMD should be limited to avoid CS, and that MR HMDs can be safe for use over extended periods of time even exceeding 25 min. Our primary recommendation for UBA is to further investigate the use of MR HMDs on RCN vessels to ensure their safe use at sea. Our future at-sea experimental campaign will help provide final input for UBA on whether MR HMDs can safely be used for underwater battlespace visualization aboard naval vessels.

References

Allcoat, D., Hatchard, T., Azmat, F., Stansfield, K., Watson, D., & von Mühlenen, A. (2021). Education in the Digital Age: Learning Experience in Virtual and Mixed Realities. *Journal of Educational Computing Research*, https:///doi/full/10.1177/0735633120985120.

Arcioni, B., Palmisano, S., Apthorp, D., & Kim, J. (2018). Postural stability predicts the likelihood of cybersickness in active HMD-based virtual reality. *Displays*. https://doi.org/10.1016/j.displa.2018.07.001.

Aslankurt, M., Aslan, L., Aksoy, A., Özdemir, M., & Dane, Ş. (2013). Laterality does not affect the depth perception, but interpupillary distance. *Journal of Ophthalmology*, 2013.

Beadle, S. C., Muth, E. R., & Pagano, C. C. (2021). Using head-mounted displays to examine adaptation and calibration under varying perturbations. *Displays*, 66, https://doi.org/10.1016/j.displa.2020.101985.

Bimberg, P., Weissker, T., & Kulik, A. (2020). On the usage of the simulator sickness questionnaire for virtual reality research. 2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW).

Cao, Z., Jerald, J., & Kopper, R. (2018). Visually-induced motion sickness reduction via static and dynamic rest frames. 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).

Casali, J. G., & Frank, L. H. (1986). Perceptual distortion and its consequences in vehicular simulation: Basic theory and incidence of simulator sickness. *Transportation Research Record*, 1059, pp. 57–65.

Champney, R., Lackey, S. J., Stanney, K., & Quinn, S. (2015). Augmented reality training of military tasks: Reactions from subject matter experts. International Conference on Virtual, Augmented and Mixed Reality,

Deepa, B., Valarmathi, A., & Benita, S. (2019). Assessment of stereo acuity levels using random dot stereo acuity chart in college students. *Journal of family medicine and primary care*, 8(12), 3850.

Eom, Y., Song, J.-S., Ahn, S.-E., Kang, S.-Y., Suh, Y.-W., Oh, J., Kim, S.-H., & Kim, H. M. (2013). Effects of interpupillary distance on stereoacuity: the Frisby Davis distance stereotest versus a 3-dimensional distance stereotest. *Japanese journal of ophthalmology*, 57(5), pp. 486–492.

Gallagher, M., & Ferrè, E. R. (2018). Cybersickness: a multisensory integration perspective. *Multisensory Research*, 31(7), pp. 645–674.

Gavgani, A. M., Walker, F. R., Hodgson, D. M., & Nalivaiko, E. (2018). A comparative study of cybersickness during exposure to virtual reality and "classic" motion sickness: are they different? *Journal of Applied Physiology*, 125(6), pp. 1670–1680.

Golding, J. F. (1998). Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain research bulletin*, 47(5), pp. 507–516.

Golding, J. F. (2006). Motion sickness susceptibility. *Autonomic Neuroscience*, 129(1–2), pp. 67–76.

Golding, J. F., Rafiq, A., & Keshavarz, B. (2021). Predicting individual susceptibility to visually induced motion sickness by questionnaire. *Frontiers in Virtual Reality*, 2, p. 3.

Gower, D., Lilienthal, M., Kennedy, R. S., & Fowlkes, J. (1988). Simulator sickness in US Army and Navy fixed-and rotary-wing flight simulators. AGARD Conference Proceedings 433. Motion Cues in Flight Simulation and Simulator Induced Sickness.

He, Z., Sui, X., Jin, G., & Cao, L. (2019). Progress in virtual reality and augmented reality based on holographic display. *Applied optics*, 58(5), pp. A74–A81.

Hemmerich, W., Keshavarz, B., & Hecht, H. (2020). Visually induced motion sickness on the horizon. *Frontiers in Virtual Reality*, 1, p. 30.

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3), pp. 1–30. doi:https://doi.org/10.1167/8.3.33.

Howarth, P. A., & Hodder, S. G. (2008). Characteristics of habituation to motion in a virtual environment. *Displays*, 29(2), pp. 117–123.

Jang, K.-M., Kwon, M., Nam, S. G., Kim, D., & Lim, H. K. (2022). Estimating objective (EEG) and subjective (SSQ) cybersickness in people with susceptibility to motion sickness. *Applied ergonomics*, 102, p. 103731.

Jasper, A., Cone, N., Meusel, C., Curtis, M., Dorneich, M. C., & Gilbert, S. B. (2020). Visually Induced Motion Sickness Susceptibility and Recovery Based on Four Mitigation Techniques. *Frontiers in Virtual Reality*, 1, p. 22.

Kemeny, A., Chardonnet, J.-R., & Colombet, F. (2020). Getting rid of cybersickness. SpringerLink.

Kemeny, A., George, P., Mérienne, F., & Colombet, F. (2017). New VR navigation techniques to reduce cybersickness. *Electronic Imaging*, 2017(3), pp. 48–53.

Kennedy, R. S., Berbaum, K. S., & Smith, M. G. (1993). Methods for correlating visual scene elements with simulator sickness incidence. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Kennedy, R. S., Drexler, J., Compton, D., Stanney, K., Lanham, D., & Harm, D. (2003). Configural scoring of simulator sickness, cybersickness and space adaptation syndrome: similarities and differences. *Virtual and adaptive environments: applications, implications, and human performance issues.* Mahwah, NJ: Lawrence Erlbaum Associates, p. 247.

Kennedy, R. S., Drexler, J., & Kennedy, R. C. (2010). Research in visually induced motion sickness. *Applied ergonomics*, 41(4), pp. 494–503. https://doi.org/10.1016/j.apergo.2009.11.006.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3), pp. 203–220.

Keshavarz, B., & Hecht, H. (2011). Axis rotation and visually induced motion sickness: the role of combined roll, pitch, and yaw motion. *Aviation, space, and environmental medicine*, 82(11), pp. 1023–1029.

Kim, H. K., Park, J., Choi, Y., & Choe, M. (2018). Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied ergonomics*, 69, pp. 66–73.

Kim, J., Luu, W., & Palmisano, S. (2020). Multisensory integration and the experience of scene instability, presence and cybersickness in virtual environments. *Computers in Human Behavior*, 113, p. 106484.

Kirollos, R., & Harriott, M. (2021). Virtual Reality to Mixed Reality Graphic Conversion in Unity. International Conference on Human-Computer Interaction.

Kirollos, R., & Jarmasz, J. (2021). *Safety considerations for the land vehicle crew training system (LVCTS): An analysis of the literature*. Defence Research and Development Canada, Scientific Report, DRDC-RDDC-2021-R101.

Kuiper, O. X., Bos, J. E., & Diels, C. (2019). Vection does not necessitate visually induced motion sickness. *Displays*, 58, pp. 82–87.

Lawson, B. (2005). Exploiting the illusion of self-motion (vection) to achieve a feeling of 'virtual acceleration' in an immersive display. Proceedings of the 11th International Conference on Human-Computer Interaction.

Lawson, B. (2014). Motion sickness symptomatology and origins. Handbook of virtual environments: Design, implementation, and applications.

Lawson, B. D., Proietti, P., Burov, O., Sjölund, P., Rodabaugh, T., Kirollos, R., & Bloch, M. (2021). *Guidelines for Mitigating Cybersickness in Virtual Reality Systems: Factors impacting cybersickness* (ch 5.2), NATO TR-HFM-MSG-323.

Lin, J.-W., Duh, H. B.-L., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. Proceedings IEEE virtual reality 2002.

Luu, W., Zangerl, B., Kalloniatis, M., & Kim, J. (2021). Effects of stereopsis on vection, presence and cybersickness in head-mounted display (HMD) virtual reality. *Scientific reports*, 11(1), pp. 1–10.

Mai, M. N., & Schlueter, M. A. (2010). The Relationship Between Pupillary Distance and Depth Perception in Humans. *Investigative Ophthalmology & Visual Science*, 51(13), pp. 4359–4359.

Merchant, W., & Kirollos, R. (2022). An Overview of Cybersickness Self Report Measures for use in DRDC Experiments. Defence Research and Development Canada, Reference Document, DRDC-RDDC-2022-D063.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12), pp. 1321–1329.

Mittelstaedt, J., Wacker, J., & Stelling, D. (2018). Effects of display type and motion control on cybersickness in a virtual bike simulator. *Displays*, 51, pp. 43–50.

Moro, C., Štromberga, Z., Raikos, A., & Stirling, A. (2017). The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical sciences education*, 10(6), pp. 549–559.

Palmisano, S., Allison, R. S., & Kim, J. (2020). Cybersickness in head-mounted displays is caused by differences in the user's virtual and physical head pose. *Frontiers in Virtual Reality*, 1, 24.

Palmisano, S., Szalla, L., & Kim, J. (2019). Monocular Viewing Protects Against Cybersickness Produced by Head Movements in the Oculus Rift. 25th ACM Symposium on Virtual Reality Software and Technology.

Porcino, T., Trevisan, D., & Clua, E. (2021). A cybersickness review: causes, strategies, and classification methods. *Journal on Interactive Systems*, 12(1), pp. 269–282.

Proietti, P., Allsop, J., Bloch, M., Bos, J. E., Burov, A., Clement, D., French, J., Kirollos, R., Lawoson, B., Leoncini, P., Liarokapis, F., Olonilua, M., Sjolund, P., Rigato, P., Rodabaugh, T., & Secci, S. (2021). Guidelines for Mitigating Cybersickness in Virtual Reality Systems [Government Report].

Prothero, J. D., & Parker, D. E. (2003). A Unified Approach to Presence and Motion Sickness. *Virtual and adaptive environments: Applications, implications, and human performance issues*, p. 47.

Rea, L. M., & Parker, R. A. (2014). *Designing and conducting survey research: A comprehensive guide*. John Wiley & Sons.

Reason, J. T. (1978). Motion sickness adaptation: a neural mismatch model. *Journal of the Royal Society of Medicine*, 71(11), pp. 819–829.

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality*, 20(2), pp. 101–125.

Rebenitsch, L., & Owen, C. (2021). Estimating cybersickness from virtual reality applications. *Virtual Reality*, 25(1), pp. 165–174.

Reed-Jones, R. J., Vallis, L. A., Reed-Jones, J. G., & Trick, L. M. (2008). The relationship between postural stability and virtual environment adaptation. *Neuroscience letters*, 435(3), pp. 204–209.

Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. *Ecological psychology*, 3(3), pp. 195–240.

Risi, D., & Palmisano, S. (2019). Effects of Postural Stability, Active Control, Exposure Duration and Repeated Exposures on HMD induced Cybersickness. *Displays,* 60, pp. 9–17.

Shafiee, D., Jafari, A. R., & Shafiee, A. A. (2014). Correlation between Interpupillary Distance and stereo acuity. *Bulletin of Environment, Pharmacology and Life Sciences,* 3(12), pp. 26–33.

Stanney, K. M., Fidopiastis, C., & Foster, L. (2020). Virtual reality is sexist: but it does not have to be. *Frontiers in Robotics and AI*, 7(4). https://doi.org/10.3389/frobt.2020.00004.

Stanney, K. M., Hale, K. S., Nahmens, I., & Kennedy, R. S. (2003). What to expect from immersive virtual environment exposure: Influences of gender, body mass index, and past experience. *Human Factors*, 45(3), pp. 504–520.

Stoffregen, T. A., & Smart Jr, L. J. (1998). Postural instability precedes motion sickness. *Brain research bulletin*, 47(5), pp. 437–448.

Teixeira, J., Miellet, S., & Palmisano, S. (2021). A Pilot Study Examining the Unexpected Vection Hypothesis of Cybersickness. Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in sport sciences*, 1(21), pp. 19–25.

Treisman, M. (1977). Motion sickness: an evolutionary hypothesis. *Science*, 197(4302), pp. 493–495.

Van Benthem, K., Cobert, H., & Zoberich, R. (2021). Human factors and ergonomics considerations when using augmented reality head mounted displays – Literature analysis report. (C21033005022). Calian.

Vrellis, I., Delimitros, M., Chalki, P., Gaintatzis, P., Bellou, I., & Mikropoulos, T. A. (2020). Seeing the unseen: user experience and technology acceptance in Augmented Reality science literacy. 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT).

Weech, S., Calderon, C. M., & Barnett-Cowan, M. (2020). Sensory down-weighting in visual-postural coupling is linked with lower cybersickness. *Frontiers in Virtual Reality*, 1, p. 10.

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. *Frontiers in psychology*, 10, p. 158.

Whittinghill, D. M., Ziegler, B., Case, T., & Moore, B. (2015). Nasum virtualis: A simple technique for reducing simulator sickness. Games Developers Conference (GDC).

XR Collaboration. (2021). *XR Glossary*. XR Ignite Inc. Retrieved June 03, 2021 from https://xrcollaboration.com/guide/a-global-resource-guide-to-xr-collaboration/xr-glossary/.

Yildirim, C. (2019). Don’t make me sick: investigating the incidence of cybersickness in commercial virtual reality headsets. *Virtual Reality*, pp. 1–9.

Annex A Demographic questionnaire

Participant Visual Perceptual Information (Completed by Experimenter)

Inter-ocular distance (in mm) _____ mm

Stereo acuity test result ______

General Information

1) Sex

M F Prefer not to say

2) Age __

Familiarity with being at sea:

Have you previously been at sea in an RCN vessel?

Y N

If you answered 'Y' to the above question, Have you experienced seasickness while at sea aboard the RCN vessel?

Y N

How many times have you been at sea aboard an RCN vessel?

When was the last time you were at sea aboard an RCN vessel?

Medical/substance Information

3) Have you recently taken any medication that is intended to reduce motion sickness (e.g., Gravol)?

Y N

Comments:

4) Are you aware of any irregularities in your vision?

Y N

Comments:

5) Are you aware of any vestibular/balance irregularities that you may have (e.g., vertigo, meniere’s disease)?

Y N

Comments:

6) Are you aware of any neurological irregularities that you may have that may affect your vestibular function (e.g. suffered from concussion; vertigo)?

Y N

Comments:

Familiarity with relevant technologies

7) I play video games Y N
8) Over the course of a week, I spend ____ hours playing video games.
9) I have used a VR or AR headset before Y N
10) Over the course of a week, I spend ____ hours using virtual reality (VR) or augmented reality (AR) headsets.
11) I have used a simulator before (e.g., flight, car, etc…) Y N

List of symbols/abbreviations/acronyms/initialisms

AR	augmented reality
CAF	Canadian Armed Forces
CS	cybersickness
D	disorientation
DND	Department of National Defence
DRDC	Defence Research Development Canada
FOV	field of view
HMD	head-mounted display
IPD	inter-pupillary distance
MR	mixed reality
MS	motion sickness
MSSQ	Motion Sickness Susceptibility Questionnaire
N	nausea
O	ocular
RCN	Royal Canadian Navy
RF	rest/reference frame
RFH	rest frame hypothesis
SR	Scientific Report
SSQ	Simulator Sickness Questionnaire
TS	total severity score
UBA	underwater battlespace awareness
VAM	vergence accommodation mismatch
VR	virtual reality
xR	extended reality

DOCUMENT CONTROL DATA

*Security markings for the title, authors, abstract and keywords must be entered when the document is sensitive

1. ORIGINATOR (Name and address of the organization preparing the document. A DRDC Centre sponsoring a contractor's report, or tasking agency, is entered in Section 8.) DRDC – Toronto Research Centre Defence Research and Development Canada 1133 Sheppard Avenue West Toronto, Ontario M3K 2C9 Canada	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) CAN UNCLASSIFIED 2b. CONTROLLED GOODS NON-CONTROLLED GOODS DMC A

3. TITLE (The document title and sub-title as indicated on the title page.)

Comparing cybersickness in two levels of virtuality in a mixed reality head-mounted display

4. AUTHORS (Last name, followed by initials – ranks, titles, etc., not to be used)

Kirollos, R.; Merchant, W.; Randall, T.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total pages, including Annexes, excluding DCD, covering and verso pages.)	6b. NO. OF REFS (Total references cited.)
October 2022	29	70

7. DOCUMENT CATEGORY (e.g., Scientific Report, Contract Report, Scientific Letter.)

Scientific Report

8. SPONSORING CENTRE (The name and address of the department project office or laboratory sponsoring the research and development.)

DRDC – Toronto Research Centre
Defence Research and Development Canada
1133 Sheppard Avenue West
Toronto, Ontario M3K 2C9
Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) APW_010	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. DRDC PUBLICATION NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC-RDDC-2022-R158	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11a. FUTURE DISTRIBUTION WITHIN CANADA (Approval for further dissemination of the document. Security classification must also be considered.)

Public release

11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Approval for further dissemination of the document. Security classification must also be considered.)

12. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Use semi-colon as a delimiter.)

mixed reality; 3D visualization; cybersickness

13. ABSTRACT (When available in the document, the French version of the abstract must be included here.)

A new underwater battlespace visualization capability is being developed for the Royal Canadian Navy. Mixed reality head-mounted displays may be promising technologies that are well-suited to provide 3-dimensional (3D) visualization of the underwater battlespace. However, cybersickness in *virtual* reality head-mounted displays is known to be a considerable health and safety risk and it remains unknown if cybersickness poses a similar challenge in *mixed* reality head-mounted displays. We present results from an experiment that assessed cybersickness in a mixed reality head-mounted display. We manipulated the amount of graphics or “virtuality” in two conditions to determine its effects on cybersickness: a “Mixed Reality” condition with low-level virtuality and a “Mixed Reality +” condition with high-level virtuality. Participants reported significantly greater cybersickness scores in the Mixed Reality + condition in later blocks than in the Mixed Reality condition. Moreover, participants reported significantly greater cybersickness scores as the experiment went on in the Mixed Reality + condition. We analyzed the impact of participant inter-pupillary distance, stereoacuity and history of motion sickness on cybersickness scores but results were not significant. Our findings led us to conclude that addition of graphic objects in mixed reality head-mounted displays will likely increase cybersickness, so their addition should be carefully considered. We also concluded that using mixed reality head-mounted displays with low-level virtuality produces negligible cybersickness for shore-based naval applications. Further investigation on board Royal Canadian Navy vessels is needed to assess the impact of ship motion on cybersickness levels.

Une nouvelle capacité de visualisation de l’espace de combat sous-marin est en cours de développement pour la Marine royale canadienne. Les visiocasques de réalité mixte pourraient être des technologies prometteuses qui conviendraient bien à la visualisation tridimensionnelle (3D) de l’espace de combat sous-marin. Cependant, l’on sait que le cybermalaise en *réalité virtuelle* est un risque considérable pour la santé et la sécurité, et l’on ignore toujours s’il pose un problème semblable dans les visiocasques en *réalité mixte*. Nous présentons les résultats d’une expérience qui a évalué le cybermalaise dans un visiocasque en réalité mixte. Nous avons manipulé le nombre de graphiques ou la « virtualité » dans deux conditions, soit en « réalité mixte » avec une virtualité faible et en « réalité mixte + » avec une virtualité élevée, afin de déterminer les effets des graphiques sur le cybermalaise. Les participants ont rapporté des scores de cybermalaise considérablement plus élevés en condition de réalité mixte + qu’en réalité mixte dans les blocs ultérieurs. De plus, les participants ont rapporté des scores de cybermalaise considérablement plus élevés au fur et à mesure de l’expérience en réalité mixte +. Nous avons analysé l’impact de l’écart interpupillaire des participants, de la stéréoacuité et des antécédents de mal des transports sur les scores de cybermalaise, mais les résultats n’étaient pas significatifs. Nos résultats nous ont amenés à conclure que l’ajout d’objets graphiques dans les visiocasques de réalité mixte augmentera probablement le cybermalaise, de sorte que leur ajout devrait être soigneusement examiné. Nous avons également conclu que l’utilisation de visiocasques de réalité mixte avec une faible virtualité provoque un cybermalaise négligeable pour les applications navales basées à terre. D’autres recherches à bord des navires de la Marine royale canadienne sont nécessaires pour évaluer l’incidence du mouvement du navire sur les degrés de cybermalaise.