The Psychometrics of Cybersickness in Augmented Reality

Claire L. Hughes*, Cali Fidopiastis, Kay M. Stanney, Peyton S. Bailey and Ernesto Ruiz

Design Interactive, Orlando, FL, United States

Augmented reality (AR) is rapidly being adopted by industry leaders and militaries around the globe. With the Defense Health Agency pushing AR as a solution to the distributed learning problem, along with AR applications being explored within primary care and operational medical settings, it is crucial for these immersive platforms to have a standardized, scientifically based paradigm on which they are designed and used. One area of particular concern is the potential for physiological maladaptation following prolonged AR exposure, which is expected to vary from that associated with virtual reality exposure. Such maladaptation is potentially driven by limitations that exist with regard to the types and extent of perceptual issues characteristic of AR head-worn displays (e.g., mismatches between visually displayed information and other senses, restricted field of view, mismatched interpupillary distance). Associated perceptual limitations can reduce training effectiveness or impose patient and/or trainee safety concerns. Thus, while AR technology has the potential to advance simulation training, there is a need to approach AR-based research—particularly that which relates to long-exposure-duration scenarios—from a bottom-up perspective, where its physiological impact is more fully understood. In the hopes of assisting this process, this study presents a comparison of cybersickness between two common forms of AR displays. Specifically, by comparing the Microsoft HoloLens, a head-worn display that has seen rapid adoption by the scientific community, with an AR Tablet–based platform within the context of long-duration AR training exposure, it will be possible to determine what differences, if any, exist between the two display platforms in terms of their physiological impact as measured via cybersickness severity and symptom profile. Results from this psychometric assessment will be used to evaluate the physiological impact of AR exposure and develop usage protocols to ensure AR is safe and effective to use for military medical training.

Keywords: augmented reality, cybersickness, virtual reality, HoloLens, AR tablet

INTRODUCTION

Over the last several years, there have been vast improvements in virtual reality (VR) and augmented reality (AR) technology, and yet, many people still report experiencing cybersickness symptoms from their use (Rebenitsch and Owen, 2016; Gavgani et al., 2017; Duzmanska et al., 2018; AR: Vovk et al., 2018; Guna et al., 2019; VR: Saredakis et al., 2020). Cybersickness is defined as the cluster of symptoms that a user experiences during or after exposure to an immersive environment (McCauley and Sharkey, 1992). It is characterized as a
physiological response to an unusual sensory stimulus, similar to motion sickness (Bouchard et al., 2007). The reported incidence and degree of intensity vary based on exposure duration and nature of virtual content and display technology; more than half of participants are expected to experience at least some degree of discomfort upon initial exposure (Lawson, 2014; Garcia-Agundez et al., 2019), although most users adapt to the environment after a few uses (Stanney et al., 2020b).

Currently the standard method for self-reporting cybersickness symptoms is the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993). This questionnaire assesses symptoms on a scale of 0 (none) to 3 (severe) and then subdivides the symptoms into three symptomatic subcategories: nausea (N), oculomotor (O), and disorientation (D). The scores for nausea relate to gastrointestinal distress (i.e., nausea, stomach awareness, salivation, and burping); scores for oculomotor relate to visual distress (i.e., eyestrain, difficulty focusing, blurred vision, and headache); and scores for disorientation relate to vestibular distress (i.e., dizziness and vertigo; Kennedy et al., 2001). These three subcategories have been used to build symptom profiles (N vs. O vs. D) associated with specific VR systems, as well as to characterize the psychometrics of cybersickness associated with VR exposure (Kennedy and Stanney, 1996; Stanney and Kennedy, 1997; Hale and Stanney, 2006; Garcia-Agundez et al., 2019; Stanney et al., 2020a).

While the typical symptom profile of D > N > O for VR has been well-characterized by previous research (Kennedy and Stanney, 1996; Stanney and Kennedy, 1997; Hale and Stanney, 2006; Garcia-Agundez et al., 2019; Stanney et al., 2020a), the same cannot be said for the adverse physiological effects of AR systems. AR devices have their own set of design challenges and potential physiological maladaptation that may differ from those associated with VR systems and even within AR systems. The symptoms may not be the same across devices; thus, the psychometrics of cybersickness in AR need to be fully characterized. The limited evidence available suggests that AR systems posses the greatest burden on the oculomotor system, specifically visual discomfort/fatigue, difficulty focusing, and headaches (Vovk et al., 2018). While studies are few, the most common symptom profile found for AR exposure is greater oculomotor disturbances (O), and at times high disorientation (D), with little nausea (N). Thus, O > D > N is the expected adverse symptom profile for AR exposure; however, further study is needed to validate that this is indeed the typical AR symptom profile. As this differs from the typical symptom profile of VR, the physiological impact of AR is expected to be distinguishable from VR systems. It is important, however, to emphasize that cybersickness is an individual problem. Each person has his/her own genetic predisposition, health history, and physical and mental attributes that influence the physiological impact of extended AR exposure. Thus, it will be important to ultimately determine an individual’s AR risk estimate, not a generalized “one size fits all” recommendation and define personalized mitigation strategies.

Physiological disturbances are expected to be compounded when an AR headset is donned for extended periods of time, as the severity of physiological maladaptation associated with VR exposure has been demonstrated to be proportional to exposure duration (Kennedy et al., 2000). Unlike VR exposure, which is oftentimes self-limiting (i.e., dropouts; Stanney et al., 1999) due to high levels of nausea and malaise, the potentially high level of oculomotor disturbance associated with AR is not expected to lead to self-cessation of exposure, as it will likely manifest as headache and eyestrain, with which people who regularly use screen-based technology are accustomed. Thus, because users will likely not self-limit exposure with AR, exposure duration could be prolonged. If AR technology poses any substantial maladaptation [e.g., prolonged adverse physiological aftereffects (AEs) such as altered visual functioning, degraded hand–eye coordination, postural instability], this could present safety risks post-exposure. It is thus of critical importance to assess the physiological impact of AR exposure and its implications to training effectiveness, patient/trainee safety, and operational advantages on the battlefield.

**TECHNOLOGY CONCERNS**

When studying physiological maladaptation possible within AR, one of the most important aspects is the technology being used for consumption. Each AR device, and even each development platform, comes with its own technology challenges that may contribute to potential for maladaptation. In general, such maladaptation is caused by some degree of mismatch between the information displayed visually within the AR display and a user’s other senses, which may be driven by low frame rate, mismatches in interpupillary distance (IPD), lag time between a user’s movement and spatial mapping of displayed information, among other factors (Fang et al., 2017). Two particularly difficult technology challenges in AR displays are vergence-accommodation conflicts and restricted field of view (FOV). Differences between a trainee’s natural depth perception and the depth planes simulated by AR may pose a particularly difficult challenge for users. Depending on the development platform used for AR generation, trainees may be forced into viewing content at specific focal distances, which may or may not match what is natively supported by the AR device, particularly those that are head-worn displays (HWDs), like the MS HoloLens. This mismatch in visual depth planes may result in a trainee perceiving depth beyond those planes that are artificially calculated and rendered (Padmanaban et al., 2017). It is likely that this process will result in physiological symptoms in the form of eyestrain, particularly related to a trainee’s natural saccadic eye movement and eye movements that occur at forced visual depth planes in an AR HWD (Fidopiastis et al., 2010). Further, when a trainee is forced to focus on depth planes optimized by the display, vergence-accommodation conflict is likely to occur. As presented depth planes approach optical infinity—which begins at approximately 6 m and is indicated by light rays being viewed as parallel by the eyes—it becomes exponentially more difficult for HWDs to replicate shifts in focus that accompany natural vision (Padmanaban et al., 2017). Such maladaptation may not be as problematic in AR-capable tablet displays.
FOV has a significant impact on the optics of HWDs (Weech et al., 2019). Humans have an FOV of ~200 degrees horizontal and ~140 degrees vertical (Mazuryk and Gervautz, 1999). Considering that human–computer interaction principles recommend a 1:1 (Buie, 1999) system of interaction as ideal, any system that constrains FOV <200 × 140 degrees will undoubtedly result in some degree of perceptual issues (Lin et al., 2002). In VR, a low FOV has been found to correlate to cybersickness (Duzmanska et al., 2018; Weech et al., 2019). While most AR HWDs have low FOV, it is unclear if the physiological impact of low FOV is as direct in AR, particularly because in AR users always have view of the real world. Specifically, as AR provides continuous viewing of real-world rest frames (e.g., walls, furniture, etc.), this may help to disambiguate virtual motion cues presented in AR HWD with vestibular cues from real-world motion or lack thereof, which should minimize cybersickness and associated adverse AEs (Chang et al., 2003). Thus, even though the HoloLens, with an FOV of 34 degrees, is significantly smaller than even the smallest VR HWD, the instantiation of virtual elements overlaid onto reality instead of directly replacing them may have a reduced maladaptive impact on users (Drascic and Milgram, 1996). Unfortunately, a wide FOV may also cause higher cybersickness levels if the FOV is paired with display stutter or similar issues (Lin et al., 2002). Thus, even as the FOV of AR displays is enlarged (Ochanji, 2020), cybersickness may persist.

There is tremendous potential to increase training efficiency with the use of AR by providing a contextually rich, embodied immersive learning environment, which allows trainees to get up to proficiency at an accelerated rate (Stamay et al., 2013; Garzón and Acevedo, 2019; Claypoole et al., 2020). However, if limitations exist regarding the type and extent of physiological maladaptation one may experience in AR-based training solutions, training effectiveness may be impeded (Lee, 2012; Fang et al., 2017) and, depending on the training task, pose safety risks should the training experience negatively impact real-world performance post-exposure (Wann et al., 2014).

**TACTICAL COMBAT CASUALTY CARE USE CASE**

In the process of updating the training curriculum for Tactical Combat Casualty Care (TCCC), the Defense Health Agency has been considering AR as a potential solution to distributed learning. TCCC is the curriculum by which the U.S. Army, Navy Corpsmen, Special Forces, Marines, and Air Force train their Combat Lifesavers (CLS). This program focuses on potentially survivable injuries that occur most often on the battlefield: the leading causes of preventable deaths being massive hemorrhage and tension pneumothorax [Bellamy, 1984; Champion et al., 2003; Butler, 2017 as cited in Kotwal et al. (2011)]. CLS provide battlefield care for these injuries while executing their unit’s mission and working to prevent further injury (National Association of Emergency Medical Technicians [NAEMT], 2018). Effective CLS training, which transfers knowledge directly and accurately to the field, is essential to decreasing preventable combat casualty deaths.

**Training Efficacy**

The potential for optimal learning training efficacy for skills such as TCCC is one of the main drivers for Department of Defense (DoD) interest in adopting AR training solutions. Currently, the standard CLS class runs over a 4-day period at most Medical Simulation and Training Centers and is required for all service members once per year. If an AR solution could increase skill retention or learning efficiency, such that the time-to-train or number of competency recertifications could be reduced, that solution would be ideal. However, if that same AR solution is causing trainees to experience cybersickness because correct design guidelines and usage protocols are not understood and in place, AR could potentially reduce training efficacy and decrease the unit’s ability to complete other required training due to adverse AEs.

One of the primary skills that all military medical providers are expected to learn is the application of a tourniquet to a casualty’s limb to treat massive hemorrhage. Tourniquets, such as the Combat Application Tourniquet favored by the military forces, are designed to stop external bleeding from a limb injury and to stabilize casualties until they can be transported to a more advanced treatment facility. Training of this skill within current AR display hardware, specifically the MS HoloLens and its presentation of depth planes, could potentially cause physiological AEs in terms of displaced hand–eye coordination that could lead to a differential of potentially several centimeters (cm) between the holographic tourniquet and real-world counterpart if the trainee were to perform the task in the real world immediately after exiting AR training. Depending on the location of the injury, this might result in the CLS provider incorrectly applying the tourniquet (i.e., negative transfer of training). In this situation, anatomic accuracy could well be the difference between life and death, as placing the tourniquet over a joint would fail to stop bleeding, likely resulting in death of a soldier.

**Safety Concerns**

Another consideration is the potential for safety risks that might arise from using AR to train TCCC medical tasks. One potential risk arises from the use of AR overlays with medical manikins. Even with recent advancements in the field, anchoring AR content to real-world objects is still challenging. A marker-based approach is generally most effective, but it requires a special marker to be aligned with the AR display device FOV at all times. Within the TCCC context, this issue could translate to one of negative training and, in turn, safety. Consider needle decompression of the chest, a treatment for tension pneumothorax. This medical intervention requires a provider to insert a needle into one of the casualty’s intercostal spaces, which is ~19.7 mm wide (Kim et al., 2014). If a three-dimensional model of a ribcage were shown to a CLS trainee as an AR overlay, and that overlay was not correctly superimposed and aligned to the medical manikin, maladaptation in hand–eye coordination could occur. The result of such negative
training might be that CLS providers who have learned to place the needle in an improper location may experience a shift in their kinesthetic position sense (Wann et al., 2014) and may perform this intervention incorrectly in the field, potentially worsening their casualty’s prospect for recovery. It is vital that when considering AR for use in training, particularly medical training, such physiological impacts on trainees be well-characterized. A study was thus conducted to characterize the psychometrics of cybersickness associated with AR exposure.

METHODS
The purpose of this study was to determine differences in cybersickness between AR HWD vs. AR Tablet training during two different exposure protocols. It was anticipated that cybersickness levels, as measured by the SSQ, would be higher in immersive HWD AR as compared to tablet-based AR and would lead to an O > D > N symptom profile in immersive AR and very low-level symptoms within the AR Tablet.

Participants
Adults aged 19–55 years (mean = 25.88, SD = 7.80), of both sexes (11 females, 23 males) participated in this study. The age and sex spread were as follows: 23 participants ≤25 years of age (7 females, 16 males); 9 participants 26–40 years of age (4 females, 5 males), and 2 participants 41–55 years of age (0 females, 2 males). This research complied with the American Psychological Association Code of Ethics and was approved by Copernicus Institutional Review Board and the Human Research Protection Office at U.S. Army Medical Research and Development Command. Informed consent was obtained from each participant, and all participants were compensated for their time in the experiment.

Equipment
The following devices were used in this study: Microsoft HoloLens 1, Samsung S5e AR-capable tablet, and a Rescue Randy medical manikin (Figure 1).

- The HoloLens 1 has 2.3-megapixel widescreen see-through holographic lenses (waveguides), a resolution of 1,280 × 720 per eye, a holographic density >2.5 K radiants, an FOV of 34 degrees with a single depth plane, and weight of 579 g (1.28 lb).
- The Samsung AR Tablet has a 10.5” WQXGA Super AMOLED display, a resolution of 2,560 × 1,600, and weight of 400 g (0.88 lb).
- A male Rescue Randy was used, which is a life-like 5’5” medical manikin with articulated joints weighing 55 lb with weight distribution according to human weight distribution chart.
Display Content

Unity game engine was used to develop immersive display content, which was focused on TCCC training. Specifically, two scenarios were developed: one focused on massive hemorrhage, and the other focused on respiration. Each scenario had timed subtasks; if a participant took the full time to complete each subtask, then the overall scenario would take 20 min to complete. However, participants could complete subtasks before the timer ran out.

- The massive hemorrhage scenario required participants to perform a tourniquet application on the manikin. During this scenario, virtual massive hemorrhage–related overlays were projected onto the physical manikin in the form of a traumatic amputation of the right leg with pulsating bleeding and pooling blood below the amputated limb.
- The respiration scenario required participants to perform a chest seal application on the manikin followed by a needle decompression of the chest after development of tension pneumothorax. During this scenario, virtual respiration related overlays were projected onto the physical manikin in the form of a left lateral open chest wound, which over time progressed to tension pneumothorax.
- Both scenarios contained training on the completion of a DD1380 Field Medical Card and the procedure for calling in a medical evacuation.

The AR display platforms allowed for physical embodiment (e.g., participants had to physically apply a tourniquet, insert a needle during chest decompression, etc.) and contextualization (e.g., scenarios placed participants in the context of the battlefield). Such physical embodiment engenders copious head and body movements, which have oftentimes been associated with motion sickness (Walker et al., 2010), whereas contextualization can add stress to training scenarios (Cohen et al., 2015).

Unity game engine was used to ensure the development of the content was as similar as possible across both AR platforms with respect to interactions and identical with respect to content. Individual differences in viewability were accounted for in the design of the TCCC experience through personalized settings, such as IPD device sizing or adjustment where necessary. A cue fidelity analysis conducted early in the effort was used to determine the optimal placement of holographic content within the real-world space. Even though Unity interacts poorly at times with FOV parameters for differing AR display types, the cue fidelity analysis allowed for control to be maintained with respect to optimal viewing of training content across the devices.

Procedure

The experiment involved five phases—prescreening, screening, pretesting, immersive exposure, and posttesting. In the prescreening phase, potential participants were referred to a weblink to take a screening survey. Any participants reporting any exclusion criteria (neurological impairments; musculoskeletal problems of the knee, ankle, shoulder, and/or elbow; loss in depth perception; \(<20/20\) corrected visual acuity; inner-ear anomalies; or history of seizures) were informed they did not qualify for participation. Participants who met prescreening eligibility went on to on-site screening, which involved informed consent and additional screening, including assessment of fitness, visual and stereo acuity, illness, alcohol, and medication consumption; participants who did not meet the criteria were excluded from the study. Participants who met screening eligibility proceeded to pretesting to complete a demographics questionnaire, have their IPD measured via a digital pupilometer, have their weight and height measured to assess body mass index, and complete surveys to assess individual demographics. During the immersive exposure phase, participants were randomized to a control (i.e., AR Tablet) or experimental group (i.e., HoloLens). Participants’ IPDs in the HoloLens group were entered into the headset software and adjusted appropriately. No participants had an IPD smaller or larger than the HoloLens range (50–80 mm). Next, marker detection was used to align virtual augmented content to the physical manikin. Once aligned, participants were exposed to the TCCC display content according to their assigned exposure protocol [three 40- (3–40 min) or six 20-min sessions (6–20 min)]. Participants commenced with their assigned starting scenario (either massive hemorrhage or respiration, counterbalanced across participants) and then alternated between the two scenarios throughout the 2-h exposure period. During the posttesting phase, SSQ total score was assessed immediately following immersive exposure (AE 0 min), and in 15-min increments for a total of 60 min (AE 15 min–AE 60 min) post-exposure, for a total of five AE measurement time periods. Participants were then debriefed, thanked, and paid for participation.

Experimental Design

The experiment was a mixed design, with 2 (exposure protocol) × 2 (display type) between factors and a 5 (AE time period) within factor. The display types between factor conditions were HoloLens headset and AR Tablet. The exposure protocol between factor conditions were 3–40 min sessions or 6–20 min sessions, both with 30-min breaks between sessions. Total AR exposure duration for each exposure protocol condition was thus 2 h. Participants were randomized to a display type (HoloLens \(n = 19\) or AR Tablet \(n = 15\)) and an exposure protocol (3–40 min sessions \(n = 13\) or 6–20 min sessions \(n = 21\)). The AE time period within factor included five post-exposure SSQ measurements, which were conducted at 0, 15, 30, 45, and 60 min post-exposure.

Dependent Measure

The dependent measure was cybersickness as measured by the SSQ (Kennedy et al., 1993) at pre-exposure baseline (BL), 0, 15, 30, 45, and 60 min post-exposure. The time component after AR exposure is critical to understanding sustained negative AEs of exposure on an individual (Stanney and Hash, 1998). To be compared appropriately, post-exposure SSQ scores were adjusted using BL pre-exposure SSQ scores. Given a total AR exposure duration of 2- and 1-h post-exposure measurement periods, participants would be expected to have “recovered” to BL SSQ levels for D, O, N, and Total SSQ scores at the conclusion of the experiment. In addition, TCCC performance measures...
Hughes et al. | Psychometrics of Cybersickness in AR

TABLE 1 | Total simulator sickness questionnaire total scores.

| Exposure protocol | Display type |   |   |   |
|-------------------|--------------|---|---|---|
|                   | HoloLens     |   |   |   |
|                   | Mean | SD | Mean | SD |   |
| 6–20 min          | Baseline   | 0.37 | 1.18 | 0.68 | 2.26 |
|                   | AE_0min     | 15.71 | 10.40 | 9.18 | 13.89 |
|                   | AE_15min    | 18.70 | 12.21 | 11.90 | 18.23 |
|                   | AE_30min    | 22.81 | 16.95 | 17.34 | 21.11 |
|                   | AE_45min    | 19.45 | 13.63 | 13.26 | 16.67 |
|                   | AE_60min    | 19.45 | 13.52 | 17.00 | 17.49 |
| 3–40 min          | Baseline   | 1.25 | 2.64 | 0.00 | 0.00 |
|                   | AE_0min     | 23.69 | 21.73 | 3.74 | 7.48 |
|                   | AE_15min    | 19.53 | 17.62 | 8.42 | 8.29 |
|                   | AE_30min    | 21.19 | 26.11 | 6.55 | 10.74 |
|                   | AE_45min    | 22.02 | 33.21 | 11.22 | 11.01 |
|                   | AE_60min    | 20.78 | 32.29 | 10.29 | 9.35 |

included number of training scenarios completed and number of correct responses.

RESULTS

As Table 1 shows, for the AR Tablet conditions, all but two post-exposure Total SSQ scores were <15.5, which puts AR Tablet in the “low” range for subjective symptomatology as compared to VR systems [see Stanney et al. (2014) for ranges]. On the other hand, the HoloLens 3–40 min condition elicited SSQ scores >20.1 but <27.9 over the duration of all AE measurement periods, which puts the HoloLens in the “medium” range for subjective symptomatology as compared to VR systems. The fact that AEs stayed elevated >30 min post-exposure places this condition in the lower 25th percentile of virtual environment systems in terms of persistence of AEs (~75% of VR systems have AEs lasting <30 min; see Stanney et al. (2014) for ranges). Stanney et al. (2014) found that if a given VR system is of medium to extreme intensity (75th percentile, with a Total SSQ score of 20.1 or higher) and is associated with persistent AEs, significant dropouts can be expected. In VR studies, dropout rates of 20% or more are common, with about 50% of attrition occurring within the first 20 min of exposure due to sickness or general malaise (Stanney et al., 1999; Reed et al., 2007). In the current study, even with SSQ > 20.1 and persistent AEs with the HoloLens 3–40 min condition, there were no dropouts. This may be due to differences in symptom profiles, which is discussed below.

A non-parametric Friedman test demonstrated that there was a significant difference for the following conditions: HoloLens, 6–20 min condition $\chi^2_r(5) = 22.75$, $p = 0.001$; HoloLens, 3–40 min condition $\chi^2_r(5) = 18.28$, $p = 0.002$; AR Tablet, 6–20 min condition $\chi^2_r(5) = 24.54$, $p = 0.001$; and AR Tablet, 3–40 min condition $\chi^2_r(5) = 11.81$, $p = 0.04$. A Wilcoxon signed-rank test with a Bonferroni correction showed that these differences were between the BL and AE Total SSQ scores ($p < 0.01$). A between-display-type evaluation was conducted for each respective exposure protocol. While Total SSQ mean scores were highest for HoloLens 3–40 min condition at the 0-min AE condition (mean = 23.69, SD = 21.73), this result was not significantly different from the AR Tablet. The confidence intervals in Figure 2 suggest that there is a potential for moderate (Total SSQ score > 15.5–20.1) to even extreme [Total SSQ scores > 33.3–53.1; see Stanney et al. (2014) for ranges] symptomatology for the HoloLens 3–40 min condition even 45 min after exposure. By 60 min post-exposure, between-participant variability in symptoms in this condition substantially contracted. The AR Tablet 6–20 min condition also has potential to reach medium cybersickness levels (Total SSQ score > 20.1–27.9) but not more extreme levels. The lesser cybersickness with 6–20 vs. 3–40 min exposures suggests that participants may have experienced a mild form of inoculation to cybersickness with more (6 vs. 3) repeat exposures due to a sensory reweighting process in which they “learned” to ignore conflicts associated with AR HWDs, such as the vergence–accommodation conflict (Stanney et al., 2020b). It is interesting that the adverse AEs persisted for >60 min post-exposure with the 3–40 min condition, which suggests that the longer exposure duration may have inhibited the inoculation process. A question-by-question analysis was conducted along with a sickness profile assessment to better understand drivers of participants’ cybersickness reports.

SSQ Subscores Symptom Profiles

Table 2 shows that the SSQ symptom profile for AR, both immersive HWD and non-immersive tablet, follows an O > D > N symptom profile, as expected. This suggests that long-duration AR exposure is associated with eyestrain, difficulty focusing, blurred vision, and headache, with lesser dizziness and vertigo, and limited nausea, stomach awareness, salivation, and burping. A one-way analysis of variance (ANOVA), $F(2,3) = 19.04$, $p = 0.049$ showed that HoloLens Oculomotor SSQ scores were significantly higher than for AR Tablet for both exposure protocols. There were no other significant differences in SSQ subscores.

Comparison of Symptom Profile Results

The SSQ questionnaire scale is interval, with a maximum value of 3 designating a severe participant response to the associated question. Weighting of subscores places more weight on Nausea (9.54) and Disorientation (13.2) than on Oculomotor (7.58) symptoms. There are also shared ratings, such as “general discomfort,” which is shared between Oculomotor and Nausea.

Figure 3 shows that fatigue and eyestrain are the dominant symptoms reported with HoloLens use regardless of exposure protocol (SSQ O subscore). Eyestrain is also dominant for the AR Tablet 3–40 min exposure condition. Oculomotor subscores were rated significantly worse for the HoloLens conditions, $X^2_{\text{HoloLens} \geq 20 \text{min}} = 23.65$, $SD = 3.10$; $X^2_{\text{HoloLens} 3–40 \text{min}} = 25.27$, $SD = 1.58$, as compared to the AR Tablet conditions, $X^2_{\text{ARTablet} \geq 20 \text{min}} = 16.26$, $SD = 3.86$; $X^2_{\text{ARTablet} 3–40 \text{min}} = 11.75$, $SD = 4.11$. The AR Tablet 3–40 min condition was comparable
Hughes et al. Psychometrics of Cybersickness in AR

FIGURE 2 | Total SSQ mean scores and 95% confidence interval.

TABLE 2 | Simulator sickness questionnaire subscores.

| Display type         | HoloLens Mean | HoloLens SD | AR Tablet Mean | AR Tablet SD |
|----------------------|---------------|-------------|----------------|--------------|
| **Exposure protocol** |               |             |                |              |
| 6 - 20min            |               |             |                |              |
| SSQ subscores        |               |             |                |              |
| Oculomotor           | 23.65*        | 3.10        | 16.26*         | 3.86         |
| Disorientation       | 14.76         | 2.11        | 12.15          | 3.86         |
| Nausea               | 9.16          | 1.45        | 6.24           | 2.41         |
| 3 - 40min            |               |             |                |              |
| SSQ subscores        |               |             |                |              |
| Oculomotor           | 25.27*        | 1.58        | 11.75*         | 4.11         |
| Disorientation       | 15.16         | 3.69        | 7.66           | 2.91         |
| Nausea               | 12.51         | 2.53        | 0.48           | 1.07         |

*p < 0.05, HoloLens significantly greater than AR Tablet.

Participants in the HoloLens 3–40 min condition reported difficulty focusing, blurred vision, and dizziness with eye closed, with some fullness of the head and nausea (SSQ D subscore; Figure 4). For the HoloLens and AR Tablet or 6–20 min conditions, participants reported low scores on most physical indicators of disorientation. Those in the AR Tablet 3–40 min condition experienced some blurred vision and fullness of the head. Figure 4 shows that participants reported difficulty focusing as their highest symptom for both HoloLens and AR Tablet 6–20 min conditions, with lesser blurred vision dizziness and fullness of the head.

Figure 5 shows that participants reported general discomfort and difficulty concentrating, as well as some stomach awareness and nausea in the HoloLens 3–40 min condition, which was the only condition to report any nausea (SSQ N subscore). The HoloLens and AR Tablet 6–20 min conditions experienced some general discomfort and difficulty concentrating.

TCCC Performance Data

Tables 3, 4 present the results of TCCC performance data, which evaluated the number of scenarios completed and number of correct responses over the cumulative exposure duration. A two-way ANOVA $F_{(1, 28)} = 17.11, p = 0.000$, showed a significant main effect of technology for both scenarios completed and correct responses. There was also a significant main effect of exposure protocol, $F_{(1, 28)} = 11.39, p = 0.002$, on scenarios completed, with more scenarios completed in the 3–40 min condition. While there was no significant interaction effect for scenarios completed or correct responses, there was a trend toward significance for correct responses, $F_{(1, 28)} = 3.57, p = 0.069$. An increased sample size may increase the significance. The results indicate that while those participants assigned to the AR Tablet condition completed more scenarios (~30% more) than those in the HoloLens group, the HoloLens group scored more correct answers than those in the AR Tablet conditions (~9% more in the 6–20 min condition; ~30% more in the 3–40 min condition). It is possible that the
FIGURE 3 | Mean oculomotor SSQ question scores with 95% CI.

FIGURE 4 | Mean disorientation SSQ question scores with 95% CI.
significantly greater oculomotor disturbances in the HoloLens condition as compared to the AR Tablet may have slowed down performance, or perhaps the novelty of the HoloLens form factor may have slowed performance. At the same time, even though less overall training content was consumed, the HoloLens condition yielded significantly better performance outcomes. The 3–40 min condition results are of particular interest, as participants in this condition who donned the HoloLens had substantially more oculomotor disturbances (mean = 25.27, SD = 1.58) as compared to the AR Tablet (mean = 11.75, SD = 4.11) but still managed to achieve substantially better performance outcomes (~30% more accurate). These results suggest that individuals may be able to overcome the adverse physiological impact of HWD AR, still concentrate on training content, and benefit from the more contextualized, embodied training afforded by this immersive form factor in order to realize deeper learning and more resilient training outcomes.

**USAGE GUIDELINES**

The SSQ results (Figures 3, 4) suggest that the burden on the visual system through eyestrain, difficulty focusing,
blurred vision, and headache is relatively pronounced in immersive HWD AR systems while performing complex, close (personal space within <2 m) training tasks such as TCCC training. Further, low levels of nausea were experienced in the HoloLens 3–40 min, but not the HoloLens 6–20 min condition (Figure 5). Given that the TCCC training scenarios were identical across these conditions, these results suggest that the adverse physiological impacts of AR exposure may be able to be moderated through usage protocols that carefully specify appropriate exposure duration in immersive AR systems; however, it is important to note that these protocols may be differentially effective based on individual differences. For example, for the immersive AR TCCC training used in this study, the 6–20 min protocol posed less of a physiological impact than the 3–40 min exposure protocol (Table 1, Figure 2). The 6–20 min condition with 30-min breaks between sessions HoloLens condition led to a Total SSQ score of 15.71 (SD = 10.4) on average immediately post-exposure, which is approaching the “low” range (25th percentile) as compared to VR systems (Stanney et al., 2014). This condition also led to strong performance outcomes in terms of correct responses (mean = 100.12, SD = 6.96). In comparison, the 3–40 min condition with 30-min breaks between sessions HoloLens condition led to a Total SSQ score of 23.69 (SD = 21.73) on average immediately post-exposure, which is firmly in the “medium” range (75th percentile) as compared to VR systems. Further, the AEs persisted in the latter condition, which would be expected to compromise post-exposure human performance and safety. However, one must also consider that the 3–40 min HoloLens condition led to strong training outcomes in terms of percent correct (mean = 105.44, SD = 20.25), and thus, participants appeared to be able to overcome the cybersickness they were experiencing and concentrate on the training content. Nonetheless, limiting exposure duration in immersive HWD AR systems to 20 min with at least 30-min breaks in between is one potential way of minimizing the adverse physiological impact of AR exposure, while still achieving strong performance outcomes. Finally, as the AR Tablet conditions had, on average, low levels of adverse symptomatology and led to substantially more training content consumed, the results suggest that complex training of longer duration may benefit from a dual technology usage protocol, where AR Tablet–based training delivers longer duration training content that is less demanding of embodied psychomotor skills and the importance of contextualization (e.g., declarative knowledge), while immersive AR headsets are used to deliver shorter-duration, fully contextualized, and embodied training experiences. An initial set of suggested AR usage guidelines thus includes the following:

- When post-exposure dexterity is important, until personalized, real-time assessment of adverse physiological effects is widely available, consider limiting exposure duration in immersive HWD AR systems to 20 min with at least 30-min breaks between exposures.
- For immersive HWD AR exposures longer than 20 min, expect dropouts and higher levels of adverse effects, such as prolonged headaches and eyestrain post-exposure, which should be measured for their severity; however, expect that trainees can overcome these adverse physiological impact and still derive substantial value from HWD AR training.
- Until adaptive AR-based training solutions are adopted, which personalize the training experience based on trainee proficiency and physiological well-being, consider adopting a dual technology usage protocol, where AR Tablet–based training delivers declarative knowledge of longer duration and immersive AR headset-based training is used to deliver shorter-duration, fully contextualized, and embodied training experiences focused on procedural and conditional (strategic) knowledge.

CONCLUSION

Immersive AR applications have the potential to significantly accelerate training expertise given the capability to present training content in a more realistic and embodied context. Understanding the potential for cybersickness and associated symptom profiles can assist in the design and development of optimal AR-based individual training protocols, such as those being developed for TCCC training. This study demonstrated an O > D > N adverse symptom profile for both immersive and tablet-based AR training systems. The oculomotor symptomatology sustained across a 60-min post-exposure assessment period for longer exposure durations (3–40 min vs. 6–20 min exposure sessions). These cybersickness indicators were, on average, of moderate to medium effects and, for oculomotor symptoms, were significantly higher and persisted longer for HoloLens conditions, as compared to AR Tablet, regardless of exposure protocol. Nonetheless, HoloLens conditions led to better performance outcomes, which suggests trainees can overcome adverse physiological impacts and still derive substantial value from HWD AR training. Thus, this preliminary research suggests that time within immersive AR training systems may need, at least initially, to be dispersed across multiple shorter exposure (~20 min) sessions with an intersession break of at least 30 min to minimize adverse symptomatology and prolonged adverse AEs, as well as foster inoculation. It was surprising to find that the persistence of adverse AEs in the long-duration immersive AR exposure condition was found to be as severe as some of the worst VR systems. Thus, even though the symptom profile for AR is loaded on oculomotor symptoms, which are less overtly incapacitating than the nausea symptoms typically associated with VR systems, the adverse symptoms can linger for long periods of time post–AR exposure, just as they do after VR exposure. More research is needed to confirm these results. Objective physiological measures of cybersickness during AR exposure, such as electrogastrography and HWD embedded eye tracking, as well as objective measures of negative adaptation effects (e.g., shifts in visual functioning, degraded hand–eye coordination, ataxia) post-exposure, should be included in future studies to quantify the extent of the effects of...
cybersickness associated with HWD AR exposure, especially on human performance.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Copernicus IRB and Human Research Protections Office at USAMRDC. The patients/participants provided their written informed consent to participate in this study. Consent has been obtained from the relevant individuals for the publication of any identifiable images in this article.

AUTHOR CONTRIBUTIONS

CH assisted with the literature review, directed the study, and was the lead author of the paper. CF led the data analytics. KS designed the experiment, contributed to, reviewed, and provided feedback on all aspects of this paper. PB and ER assisted with the literature review, collected the data, reviewed, and provided feedback on the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This research and development project/program/initiative was conducted by Design Interactive, Inc and is made possible by a contract that was awarded and administered by the U.S. Army Medical Research & Development Command (USAMRDC) and the Joint Program Committee–JPC-1 at Fort Detrick, MD under Contract Number: MTEC-W81XWH1900005.

ACKNOWLEDGMENTS

This paper was dedicated to Dr. Robert S. Kennedy, who was a world-leading expert in motion sickness and a dedicated, generous, and humble mentor, giving endlessly to those fortunate enough to be honored by his thoughtful tutelage. Bob was a Gentleman Scientist, reminiscent of other luminaries for whom the pursuit and sharing of knowledge was first and foremost their highest calling. Additionally, the authors would like to thank Darrin Frye, Frank Karluk, and the rest of the JPC-1 team for their guidance throughout this effort.

REFERENCES

Bellamy, R. F. (1984). The causes of death in conventional land warfare: implications for combat casualty care research. Mil. Med. 149, 55–62. doi: 10.1093/milmed/149.2.55
Bouchard, S., Robillard, G., and Renaud, P. (2007). Revising the factor structure of the simulator sickness questionnaire. Ann. Rev. CyberTherapy Telemed. 5, 128–137. doi: 10.1037/1763-0606-000
Buie, E. (1999). HCI standards: a mixed blessing. Interactions 6, 36–42. doi: 10.1145/296165.296177
Butler, F. K. Jr. (2017). Tactical combat casualty care: beginnings. Wilderness Environ. Med. 28, 12–17. doi: 10.1016/j.wem.2016.12.004
Champion, H. R., Bellamy, R. F., Roberts, C. P., and Leppaniemi, A. (2003). A profile of combat injury. J. Trauma 54, 13–19. doi: 10.1097/01.TA.0000057151.02906.27
Chang, E., Hwang, L., Jeon, H., Chun, Y., Kim, H. T., and Park, C. (2003). “Effects of rest frames on cybersickness and oscillatory brain activity,” in International Winter Workshop on Brain Computer Interface (Los Alamitos, CA: IEEE Computer Society Press), 62–64.
Claypool, V. L., Padron, C. K., Sanchez, S. A., Hodges, C. A., Stanney, K. M., and Perez, R. S. (2020). “Leveraging emerging technologies for shipboard training and sailor support: recent advancements and current operational challenges,” in American Society of Naval Engineers (ASNE) 2020 Fleet Maintenance and Modernization Symposium (FMMS). Virtual Symposium Sept 15-16 (San Diego, CA).
Cohen, I., Brinkman, W., and Neerincx, M. A. (2015). Modelling environmental and cognitive factors to predict performance in a stressful training scenario on a naval ship simulator. Cogno. Technol. Work 17, 503–519. doi: 10.1007/s10111-015-0325-3
Draslic, D., and Milgrom, P. (1996). “Perceptual issues in augmented reality,” in Proceedings of Stereoscopic and Virtual Reality Systems III, eds S. S. Fisher, M. T. Bolas, and J. O. Merritt (Bellingham WA: SPIE), 2653. doi: 10.1117/12.237425
Hughes et al. Psychometrics of Cybersickness in AR
Kennedy, R., and Stanney, K. (1996). Postural instability induced by virtual reality exposure: development of a certification protocol. Int. J. Hum. Comput. Interact. 8, 25–47. doi: 10.1080/10447318.1996.9521639

Kennedy, R. S., Stanney, K. M., and Dunlap, W. F. (2000). Duration and exposure to virtual environments: sickness curves during and across sessions. Presence 9, 463–472. doi: 10.1162/10447318.2000.9566952

Kim, Y. S., Park, M. J., Rhim, H., Lee, M. W., and Lim, H. K. (2014). Sonographic analysis of the intercostal spaces for the application of high-intensity focused ultrasound therapy to the liver. AJR 203, 201–208. doi: 10.2214/AJR.13.11744

Kotwal, R. S., Montgomery, H. R., Champion, H. R., Butler, F. K., Mabry, R. L., et al. (2011). Eliminating preventable death on the battlefield. Arch. Surg. 146, 1350–1358. doi: 10.1001/archsurg.2011.213

Lawson, B. D. (2014). “Motion sickness symptomatology and origins,” in Kotwal, R. S., Montgomery, H. R., Kotwal, B. M., Champion, H. R., Butler, F. K., McCauley, M. E., and Sharkey, T. J. (1992). Cybersickness: perception, TechTrends

Lee, K. (2012). Augmented reality in education and training. TechTrends 56, 13–21. doi: 10.1007/s11528-012-0559-3

Lin, J., Duh, H., Parker, D., Abi-Rached, H., and Furness, T. (2002). “Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment,” in Proceedings - Virtual Reality Annual International Symposium (Seattle, Washington DC), 164–171. doi: 10.1109/VR.2002.996519

Mazuryk, T., and Gervautz, M. (1999). Virtual Reality: History, Applications, Technology, and Future. Vienna: Institute of Computer Graphics, Vienna University of Technology. Available online at: https://www.cg.tuwien.ac.at/research/publications/1996/mazuryk-1996-VRH/TR-186-2-96-06Paper.pdf (accessed March 9, 2020).

McCaeley, M. E., and Sharkey, T. J. (1992). Cybersickness: perception of self-motion in virtual environments. Presence 1, 311–318. doi: 10.1162/pres.1992.1.3.311

National Association of Emergency Medical Technicians [NAEMT], (2018). Tactical Combat Casualty Care for medical personnel August 2018. Available online at: www.naemt.org/docs/default-source/education-documents/tccc/tccmp/presentations/1-introduction-to-tccc.pptx?sfsrvsn=-3953fc892_2 (accessed March 9, 2020).

Ochanji, S. (2020). Ostendo unveils new AR display with 150 degrees field of view. VR Times. Available online at: https://virtualrealitytimes.com/2020/07/19/ostendo-unveils-new-ar-display-with-150-degrees-field-of-view/ (accessed September 3, 2020).

Padmanaban, N., Konrad, R., Stramer, T., Cooper, E. A., and Wetzstein, G. (2017). Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. Proc. Natl. Acad. Sci. U.S.A. 114, 2183–2188. doi: 10.1073/pnas.1617251114

Rebenitsch, L., and Owen, C. (2016). Review on cybersickness in applications and visual displays. Virtual Real. 20, 101–125. doi: 10.1007/s10055-016-0285-9

Reed, N., Diels, C., and Parkes, A. M. (2007). “Simulator sickness management: enhanced familiarisation and screening process,” in Proceedings of the First International Symposium on Visually Induced Motion Sickness, Fatigue, and Photosensitive Epileptic Seizures (VIMS2007), eds R. H. Y. So, R. T. F. Cheung, E. H. C. Chow, J. T. T. Li, and A. K. C. Lam (Hong Kong: University of Hong Kong), 156–162.

Saredakis, D., Szpak, A., Birckhead, B., Keage, H., Rizzo, A., and Loetscher, T. (2020). Factors associated with virtual reality sickness in head mounted displays: a systematic review and meta-analysis. Front. Hum. Neurosci. 14.90. doi: 10.3389/fnhum.2020.00906

Stanney, K., Carroll, M., Champrey, R., DeVore, L., and Hale, K. (2013). “Virtual environment training design: pros, cons and tailored solutions,” in Fundamental Issues in Defence Training and Simulation, eds C. Best, G. Galanis, J. Kerry, and R. Sotilire (Aldershot, UK: Ashgate), chapter 15.

Stanney, K. M., Fidopiastis, C., and Foster, L. (2020a). Virtual reality is sexist: but it does not have to be. Front. Robot. AI 7:4. doi: 10.3389/frobt.2020.00004

Stanney, K. M., and Hash, P. (1998). “Locus of user-initiated control in virtual environments: influences on cybersickness.” Presence 7, 447–459. doi: 10.1162/10547498565848

Stanney, K. M., and Kennedy, R. S. (1997). The psychometrics of cybersickness. Commun. ACM 40, 66–68. doi: 10.1145/257874.257889

Stanney, K. M., Kennedy, R. S., and Hale, K. S. (2014). “Virtual environments usage protocols,” in Handbook of Virtual Environments: Design, Implementation, and Applications, 2nd Edn, eds K. S. Hale and K. M. Stanney (New York, NY: CRC Press), 797–809. doi: 10.1201/b17360-38

Stanney, K. M., Lanham, D. S., Kennedy, R. S., and Breaux, R. (1999). Virtual environment exposure drop-out thresholds. Proc. Human Factors Ergon. Soc. Ann. Meet. 43, 1223–1227. doi: 10.1177/1541931299390430212

Stanney, K. M., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., et al. (2020b). Identifying causes of and solutions for cybersickness: reformulation of a research and development agenda. Int. J. Hum. Comput. Interact. 36, 1783–1803. doi: 10.1080/10447318.2020.1828535

Vovk, A., Wild, F., Guest, W., and Kuula, T. (2018). “Simulator sickness in augmented reality training using the Microsoft HoloLens,” in Proceedings of the CHI Conference on Human Factors in Computing Systems (New York, NY: Association for Computing Machinery), 1–9. doi: 10.1145/3173574.3173783

Walker, A. D., Muth, E. R., Switzer, F. S., and Hoover, A. (2010). Head movements and simulator sickness generated by a virtual environment. Aviat. Space Environ. Med. 81, 929–934. doi: 10.3357/ASEM.2735.2010

Wann, J. P., White, A. D., Willkie, R. M., Culmer, P. R., Lodge, P. A., and Mon-Williams, M. (2014). “Measurement of visual aftereffects following virtual environment exposure: implications for minimally invasive surgery;” in Handbook of Virtual Environments: Design, Implementation, and Applications, 2nd Edn, eds K. S. Hale and K. M. Stanney (New York, NY: CRC Press), 811–833. doi: 10.1201/b17360-39

Weech, S., Kenny, S., and Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. Front. Psychol. 10:158. doi: 10.3389/fpsyg.2019.00158

Disclaimer: The views, opinions and/or findings contained in this research/presentation/publication are those of the authors/company and do not necessarily reflect the views of the Department of Defense and should not be construed as an official DoD/Army position, policy or decision unless so designated by other documentation. No official endorsement should be made. Reference here to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government.

Conflict of Interest: CH, CF, KS, PB, and ER were employed by the company Design Interactive.

Copyright © 2020 Hughes, Fidopiastis, Stanney, Bailey and Ruiz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.