Health and Safety of VR Use by Children in an Educational Use Case

Robert Rauschenberger, Ph.D.*
Phoenix User Research Center
Exponent, Inc.

Brandon Barakat, Ph.D.†
Phoenix User Research Center
Exponent, Inc.

ABSTRACT

The present study examined the potential health and safety effects of short-term virtual reality (VR) use by children in an educational use case scenario (that is, relatively brief episodes of use across a limited number of sequential days), such as how VR may be used in the classroom or at a museum. Ophthalmological, vestibular functioning, balance, hand-eye coordination, 3D spatial representation, and subjective comfort effects were assessed using a variety of optometric, psychophysical, and self-report measures. Thirty child participants (ages 10 to 12 years) were immersed in VR for 30 minutes daily across five consecutive days of use. Measurements were taken prior to the onset of VR use (baseline), at the end of the fifth day of VR use (to assess potential acute effects), and 24 hours after the fifth day of VR use (to assess potential longer-lasting effects).

There were no statistically significant adverse effects found, with the exception of slightly elevated scores on a self-reported measure of subjective comfort, which, however, were below the range of scores reported in past research as being indicative of subject discomfort. In other words, the current study found no empirical evidence that short-term use of VR in an educational use setting by children ages 10 through 12 years is associated with any adverse visual, spatial representational, or balance aftereffects, or that it causes undue nausea, oculomotor discomfort, or disorientation. The present study does not address longer-term use or potential psychological effects of different VR content.

Keywords: Virtual reality, children, education, health and safety, SSQ, optometry, balance, hand-eye coordination, spatial representation.

Index Terms: [Human-centered computing—Interaction Paradigms]:Virtual Reality; [User Characteristics]:Age—Children

1 INTRODUCTION

There is much speculation in the popular press about the potential health and safety consequences of virtual reality (VR) use [1, 2]. At the same time, there is little published scientific data on this topic, at least with regard to the current generation of consumer-grade VR devices, which have outpaced their predecessor devices of decades past, in terms of quality, comfort, and perceptual realism. With the continued ascent of VR as an emerging consumer technology with potential widespread applications, it becomes increasingly important to bring to bear scientific data on the speculation that exists about whether VR can be used safely and without undue adverse effects in specific settings. Indeed, VR is already being used in a variety of settings and by a variety of user types, including, for example, as a research tool to study various behaviors in children [3–5]. Another setting of potential widespread interest is an educational use case [6]. Teachers may want to employ VR to transport students into a specific geographic region or a particular historical moment with an unprecedented degree of realism and sense of presence. Similarly, museums may wish to augment their physical displays by curating virtual exhibits, presented in VR. Students may return to the museum on multiple sequential days during a field trip or as part of a museum-based class; or teachers may want to spend several days on the same topic, using VR on multiple days.

The aim of the present study, therefore, was to examine the potential health and safety effects of short-term VR usage on children, within an educational use case scenario. Specifically, participants spent relatively brief periods of time (30 minutes) in VR, across a total of five consecutive days. Because it is unlikely that students will spend such a span of time in VR, the present study tests a somewhat extreme version of the educational use case, to "stress test" the bodily systems involved. The study also did not contend itself to examine a single possible aftereffect, as other studies have, but rather essayed to span a spectrum of potential adverse effects. These included: potential ophthalmological effects [7], potential effects on balance and vestibular functioning [8], potential effects on the representation of three-dimensional space and hand-eye coordination [9], and possible effects on subjective comfort and well-being [10].

Much of the speculation about potential health and safety consequences of VR use stems from studies documenting adverse effects from the use of early simulators [11], which, in some ways, are large-scale versions of modern VR headsets. Long before the advent of consumer VR products, simulators have been evaluated for potential adverse effects, going as far back as at least the 1950s, when flight simulators were first used to train military pilots [12]. The study of simulator aftereffects entered a new era with the introduction of the Simulator Sickness Questionnaire (SSQ) in the 1990s [13], which permitted a more systematic study of the effects of simulation on study participants. Given the long history of the SSQ and its ubiquitous use in the study of VR aftereffects, the SSQ was also employed in the present study.

The other measures collected in the present study were adopted from relevant areas of scientific literature and are measures of variables that have been implicated as potential adverse effects of VR use. It has been suggested, for example, that the vergence-accommodation conflict engendered by the presentation of visual content with varying amounts of simulated depth at a fixed distance from the eyes can lead to measurable effects on visual functioning [14, 15]. Similarly, it has been suggested that the vestibular-ocular disconnect created by the dissociation of bodily motion and simulated motion can result in nausea and imbalance [16]. Lastly, it is known from very early studies of perceptual adaptation that a pervasive remapping of visual and/or proprioceptive inputs can require re-adaptation when returning to normal inputs, which can result in certain sensorimotor aftereffects [17].

In the present study, each of these potential effects is assayed by at least one measure, as illustrated in Figure 1. Indeed, most of the effects are probed by multiple measures to provide converging data. All of the measures employed have been used in previous, peer-reviewed scientific studies [13, 18–22].

*e-mail: rrauschenberger@exponent.com
†e-mail: bbarakat@exponent.com

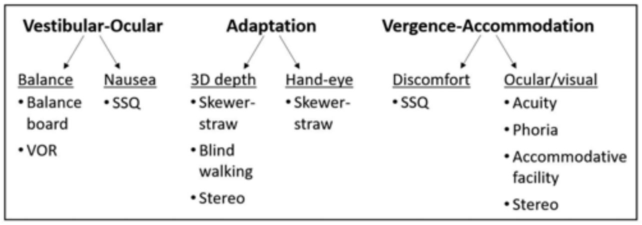

Figure 1: Potential aftereffects of VR use and measures employed in the present study to evaluate the presence of these possible effects.

2 Methods

2.1 Participants

Thirty naïve participants (15 female, 15 male), ages 10 through 12 years (mean: 11.0), were recruited for participation in the study. This age range was selected for two reasons. First, the majority of consumer VR devices are labeled as being intended for users 12 or 13 years old and up; therefore, the current study's age group allows for inferences to be made about a small range of ages that is just below the age range that is currently accepted to be safe by most manufacturers. Second, according to indications in the literature, there exists an interest in using VR for educational purposes with students in the 10 to 12 year age range [6].

All participants were compensated $400 for their participation. Prior to their participation, participants read and signed an assent form, and their accompanying parent or legal guardian signed a consent form on their behalf. Participants and their parent or legal guardian were reminded that participation in the study was completely voluntary. All procedures, stimulus materials, and participant forms were approved by an Institutional Review Board (IRB) prior to the commencement of the study.

2.2 Stimuli and Design

In order to obtain a reasonable sampling of the current consumer VR landscape, participants were asked to view content in one of five VR devices: the Lenovo Mirage Solo with Google Daydream [23], the HTC Vive [24], the Oculus Rift [25], the Sony PlayStation VR [26], or the Samsung Gear VR [27]. The focus of the study was not any manufacturer's specific product but the use of VR per se by children in an educational context and, therefore, use of the individual manufacturer devices was proportioned equally across participants. In other words, all fives devices were used by an equal proportion of participants (i.e., six participants per device), and each participant used only one of the five devices throughout his/her participation. The technical specifications for each of these VR devices (e.g., resolution, refresh rate, field of view, etc.) can be found on their respective product web pages [23–27]. Each device's default graphics settings were used throughout the study.

Given that digital educational content has the potential to comprise a wide variety of formats and genres, ranging from passive-viewing experiences to interactive game-based learning experiences [6, 28], the VR content used in the current study was curated to be representative of this varied landscape. Specifically, each VR device was intentionally loaded with a collection of content that contained varying levels of both passive-viewing and interactive elements. Although certain pieces of content were not necessarily educational, they were included to ensure each device contained content that covered the aforementioned range of both passive-viewing and interactive game-based elements. Furthermore, all devices contained content that was age-appropriate and that did not require participants to walk around.

Although an attempt was made to have similar content across all five VR devices, not all content was available on all devices. Specifically, depending upon the content that was available on the specific headset they used, participants were offered a choice of at least five of the following VR content: Beat Saber, Carpe Lucem, Daydream tutorial, Discovery VR, Invasion!, Job Simulator, Moss, Oculus Prologue, Smashbox (Arena), Star Chart (VR), Steam VR Tutorial, Tethered, Titans of Space, Virtual Virtual Reality, and Virush. All participants were instructed to try all of the content that was available on their VR device at least once, but they were otherwise free to choose which content they wanted to play throughout their VR sessions. Furthermore, all participants were instructed to remain seated throughout their VR usage and they were monitored at all times by a study moderator.

All VR sessions and measurements were conducted within usability labs at Exponent's Phoenix User Research Center (PURC) [29]. PURC comprises an interdisciplinary team of user researchers and large-scale usability labs suitable to observations of participants engaged in VR use, as well as a full optometry lab, balance measurement capabilities, motion tracking cameras, and other equipment appropriate for the study of VR.

2.2.1 Dependent measures

Potential aftereffects of VR use were assessed using four types of measures: ophthalmological, vestibular/balance, psychophysical, and subjective comfort. All measures (with the exception of the SSQ) were administered before participants entered VR on the first day (baseline), at the end of the fifth day of VR use (to assess potential acute effects), and 24 hours after the fifth day of VR use (to assess potential longer-lasting effects). The SSQ was administered only after VR use, to mitigate any demand characteristics (see below). The measurements that were recorded on the sixth day, which did not involve the use of VR, were used to determine whether any effects – should they have been determined to be statistically significant in an adverse direction on the fifth day – still persisted 24 hours following the cessation of VR use.

2.2.2 Ophthalmological measures

The ophthalmological measures included acuity, phoria (AC/A ratio), accommodative facility, and stereo perception. All measurements were done by a professional practicing optometrist. Acuity was measured for both near and far viewing, using a Reichert phoropter, a Jaeger reading card (for near acuity), and a display with Snellen letters (for far acuity), such as is used for eye exams when determining a prescription for corrective lenses. Phoria was measured using an alternating cover test with prism neutralization, in which one eye is covered and any deviation in the uncovered eye is neutralized using a prism bar [30]. Measures were taken at near (40 cm) as well as far (clinical infinity) distance. Accommodative facility was measured by use of a "flipper." The number of flips from +2D to -2D (and vice versa) was recorded as participants read out letters printed in 2-point font once they appeared visually clear to them [18]. This task was performed for 30 seconds and the resulting number of flips was then multiplied by two to calculate the corresponding standard per-minute rate. Finally, stereo perception was measured using a Stereo Optical stereo vision test and polarized glasses [31]. Specifically, the test consisted of nine sets of four concentric-circle targets. In each set, the participant indicated which of the four targets "popped out" from the page. The disparity of the targets ranged from 800 to 40 seconds of arc, with 40 arc seconds being the most difficult to perceive.

2.2.3 Balance and vestibular functioning measures

Balance was recorded using an Otometrics ICS Balance Platform under four conditions: eyes open/no cushion, eyes closed/no cushion, eyes open/cushion, and eyes closed/cushion [19]. That is, at each measurement time point, each participant's postural stability was measured four times, including with their eyes open and closed, and with and without the participant standing on a cushion that was placed on top of the balance platform. During each measurement, participants were instructed to stand as still as possible, with their arms across their chest, while their postural stability was measured for 20 seconds. Of the available metrics, anterior-posterior sway, lateral sway, and total sway area were exported from the device for analysis. Specifically, anterior-posterior sway provided a measure of variability in a participant's stance along the front-back axis, lateral sway provided a measure of variability in a participant's stance along the left-right axis, and total sway area provided a measure of the two-dimensional area of space that was covered by deviations in a participant's center of gravity.

Balance measurements were recorded at four time points, with the first time point being on the first day prior to VR use (baseline). Following VR use on the fifth day, balance was measured twice, including immediately after VR use (day 5, post VR 1) and again approximately 20 minutes after VR use (day 5, post VR 2), to determine if any imbalance detected immediately after VR use might normalize after a slight delay. Balance measurements were then recorded again 24 hours following VR use (day 6, no VR), to assess whether potential adverse effects on balance that were observed on the fifth day persisted into the following day.

Vestibular functioning, in turn, was assessed using the "head impulse test" [20], which evaluates the intactness of the Vestibulo-Ocular Reflex (VOR). To evaluate the intactness of the VOR, the participant was asked to fixate on the tip of the researcher's nose while the researcher gently rotated the participant's head clockwise and counter-clockwise. The researcher then briskly rotated the participant's head and monitored for uninterrupted fixation of the tip of the nose, which indicates an intact VOR. This procedure was performed twice at each time point, with the brisk rotation being made once in each direction (i.e., clockwise and counterclockwise), to assess any potential unilateral effects on the VOR.

2.2.4 Psychophysical measures

Three-dimensional spatial representation and hand-eye coordination were evaluated using two measures: a near-depth perception task (the skewer-straw task) [21] and a far-depth perception task (the blind walking task) [22].

In the skewer-straw task, participants inserted 25 skewers into randomly oriented straws, one at a time, and the total time to complete the task was recorded [21]. Performance in this task depends upon a combination of dexterity, stereo vision, and hand-eye coordination. On each repetition of the task, a different side of the apparatus was oriented toward the participant, to prevent improvements in performance across trials that were due to spatial memory. Prior to data collection proper, all participants completed one practice round, because past research has indicated that there is an initial acclimation to the task that occurs after one practice trial [21].

In the blind walking task, participants were asked to walk while blindfolded to one of two targets: a near target placed at 2 meters from the starting line and a far target placed at 4 meters from the starting line [22]. Across trials, participants alternated walking to the near and the far target from the starting line. They made five attempts at each target. On each trial, the participants' error in target estimation (i.e., the difference between the distance to the target and the distance walked) was recorded by the researcher.

On each trial, participants were allowed to look at the target location while standing at the starting line, but then walked to the target blindfolded and did not receive feedback even after indicating that they had reached the designated target. The reason participants were deprived of feedback is that, even without an accurate three-dimensional representation of the blind-walking space, participants could have used the magnitude and direction of their errors to minimize the distance to each target. Therefore, by depriving participants of feedback on their accuracy, the task provides a measure of potential distortions in far-distance depth judgments [22].

2.2.5 Subjective comfort measures

Subjective comfort was evaluated using the SSQ, which comprises three subscales: nausea (SSQ-N), oculomotor discomfort (SSQ-O), and disorientation (SSQ-D) [13]. The SSQ contains 16 questions about the respondent's currently-experienced symptoms, such as increased salivation, sweating, eye strain, etc. Respondents indicate the level at which they are experiencing each of these symptoms on a four-point scale ranging from "None" to "Severe." These numeric responses are then summed and weighted separately for each of the subscales. Scores on the individual subscales can, theoretically, range from zero to 200.34 (SSQ-N), 159.18 (SSQ-O), and 292.32 (SSQ-D). As noted in the original publication that introduced the SSQ, there is no literal meaning to the resulting subscale scores, but rather they can be used for making comparisons between measurements [13].

In past research, the SSQ has been shown to be vulnerable to demand characteristics – i.e., the simple act of administering the questionnaire both before and after VR exposure gives participants the impression that they are expected to report an effect of VR, and this impression alone can influence their responses on the questionnaire, separate from any potential effect of VR itself [32]. In the present study, therefore, SSQ responses were collected only after VR use, to mitigate any demand characteristics.

3 Results

All data were analyzed using repeated measures analysis of variance (ANOVA) as well as planned comparisons that specifically compared the baseline measurements to the measurements taken immediately following VR use on day 5, and that compared the baseline measurements to the measurements taken 24 hours following VR use on day 6 (with the exception of the SSQ data, for which no baseline values were recorded, to mitigate demand characteristics – see above). In each figure, the error bars represent 95% within-subject confidence intervals [33].

3.1 Ophthalmological measures

Participants' measured acuity did not differ at the three measurement time points: day 1 (baseline), day 5 (post VR use), and day 6 (24 h post); $F < 1$. The average Snellen-equivalent values recorded were 20.3, 21.6, and 20.3, respectively, for the three time points. Similarly, participants did not exhibit a statistically significant change in phoria across the three measurement time points; $F < 1$.

By contrast, the values recorded for accommodative facility did differ across the three measurement time points; $F(2, 58) = 5.90$, $p < .05$. However, this represented an improvement in performance across time points, with the number flips numerically increasing from day 1 to day 5 and then slightly again on day 6. The average number of flips achieved in 60 seconds was 10.4, 11.3, and 11.7, respectively, for the three measurement time points.

Lastly, stereo perception did not change across the three measurement time points; $F(2, 58) = 2.15$, n.s. Participants were able to achieve a stereo acuity of 41.0, 42.8, and 40.7 seconds of arc, respectively, for the three measurement time points (i.e., nearly perfect stereo vision scores, on average).

In sum, there were no notable changes in acuity, phoria, accommodative facility, or stereo perception subsequent to being immersed in VR for 30 minutes daily across five days.

3.2 Balance and vestibular functioning measures

All participants exhibited a normal VOR in the head impulse test at all three time points: day 1 (baseline), day 5 (post VR use), and day 6 (24 h post). There was, moreover, no indication of systematic or pervasive changes in balance, as illustrated in Figure 2. Specifically, of the 36 planned comparisons that were made – i.e., three time point comparisons (baseline compared to the three post-VR measurements) for each combination of three balance metrics

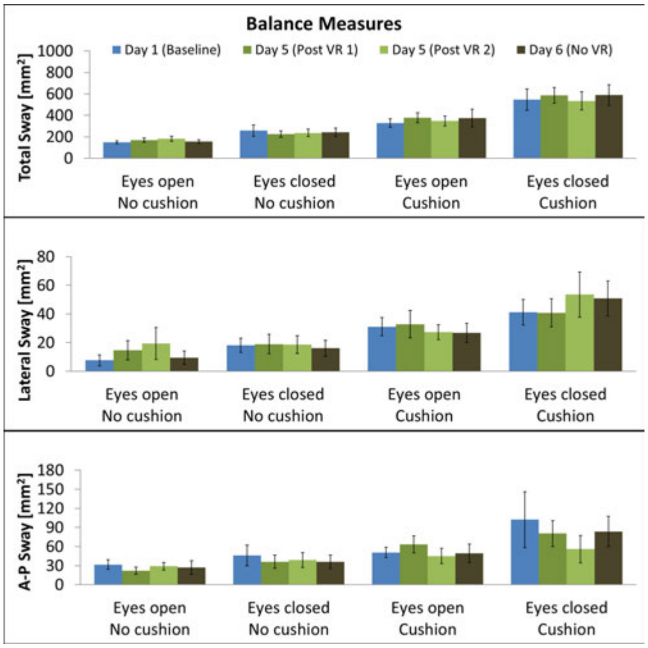

Figure 2: Three measures of balance (Total Sway area, Lateral Sway, Anterior-Posterior Sway), for four different conditions (Eye open/No cushion, Eyes closed/No cushion, Eyes open/Cushion, Eyes closed/Cushion), averaged across participants, for the first day (baseline), the fifth day, and the sixth day of the study. No VR immersion occurred on the sixth day. Two post measures (Post VR 1, Post VR 2) were taken 20 minutes apart on the fifth day, to evaluate whether any potential effects dissipated by the second measurement. Error bars represent 95% confidence intervals.

(anterior-posterior sway, lateral sway, and total sway area) and four balance conditions (eyes open/no cushion, eyes closed/no cushion, eyes open/cushion, and eyes closed/cushion) – only two of the comparisons were found to be statistically significant, and these two effects occurred in opposite directions of each other. These were a decrease in anterior-posterior sway in the first post-VR measurement on day 5 compared to baseline, $t(29) = 2.33$, $p < .05$, and an increase in total sway area in the second post-VR measurement on day 5 compared to baseline, $t(29) = 2.07$, $p < .05$, both of which occurred in the eyes open/no cushion condition. Furthermore, when examining the balance data in aggregate, it is not clear that any overarching pattern of effects can be discerned across time points, balance conditions, or balance metrics.

In sum, then, these data give no indication that engaging in daily use of VR for 30 minutes over the course of five days adversely affects the balance or vestibular functioning of children ages 10 through 12 years.

3.3 Psychophysical measures

As can be seen from Figure 3, the completion times for the skewer-straw task were numerically elevated following five days of VR immersion – but not statistically significantly elevated; $F(2, 58) = 2.23$, n.s. Planned comparisons revealed that neither the comparison between baseline and day 5 nor the comparison between baseline and day 6 reached statistical significance; $t(29) = 1.99$, n.s. and $t(29) = 1.84$, n.s., respectively.

Figure 4 shows the results from the blind walking task. The interpretation of these results is complicated by the fact that, according to past research, repetition of the blind walking task results in progressive increases in the distance participants walk, independent of any manipulation that could otherwise affect the accuracy of their

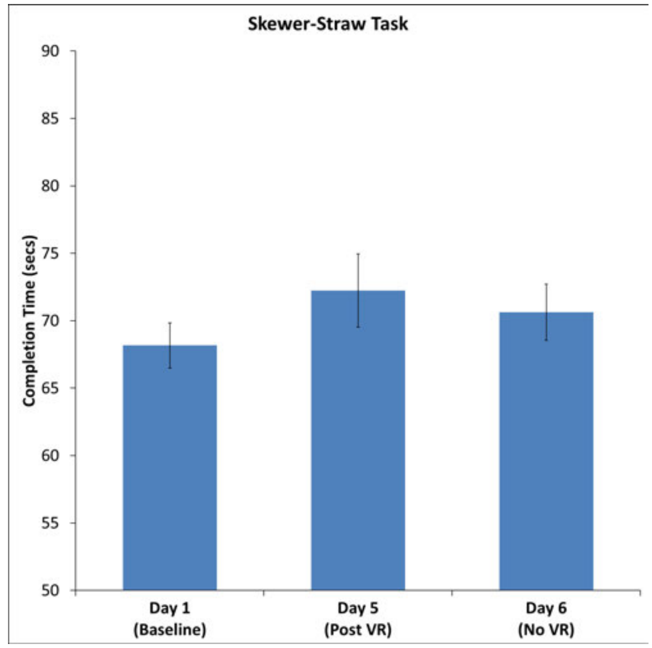

Figure 3: Average completion times for the skewer-straw task on days 1, 5, and 6 of the study. No VR immersion occurred on day 6. Error bars represent 95% confidence intervals.

performance (i.e., participants simply walk farther and farther across consecutive trials) [34]. Among the potential reasons given for why participants walk increasingly greater distances in the blind walking task are that they feel more confident walking blindly after several repeated attempts; they increase their stride length; and they increase their walking pace [34].

Indeed, when the current study's blind walking data were examined on a trial-by-trial basis, this artifact of task repetition was clearly evident, with participants' walking distances increasing monotonically across consecutive trials, at all time points, and for both the near and far targets. To verify whether this artifact was present, the five baseline trials were examined alone (i.e., trials that were free of any potential influence of VR), and the data were found to have a statistically significant fit to a linear function for both the near and far target; $t(29) = 3.50$, $p < .05$ and $t(29) = 3.27$, $p < .05$, respectively. Therefore, in an effort to mitigate the influence of this artifact on the dataset, the blind walking data were analyzed using only the first trial at each time point, for each target location.

As can be seen in Figure 4, there was no change in participants' blind walking distances between baseline and day 5, for either the near or the far target; $t(29) = 1.19$, n.s. and $t(29) = 0.62$, n.s., respectively. There was a statistically significant change between baseline and day 6, for both the near and the far target; $t(29) = 4.85$, $p < .05$ and $t(29) = 3.63$, $p < .05$, respectively. Although, technically, it is ambiguous as to whether this difference is attributable to an effect of VR or to the task repetition artifact, it is much more likely that the increase on day 6 reflects the cumulative effect of the task repetition artifact.

Firstly, for there to have been an effect of VR on day 6, but not on day 5, it would constitute a sort of delayed-onset perceptual aftereffect, which would be entirely contrary to what is known about the time course and underlying mechanisms of perceptual aftereffects [35]. Such perceptual aftereffects are strongest immediately upon transitioning from the perceptually distorted environment in which sensory adaptation occurred, back to the undistorted physical world. Real time perceptual experience with the latter leads to re-adaptation

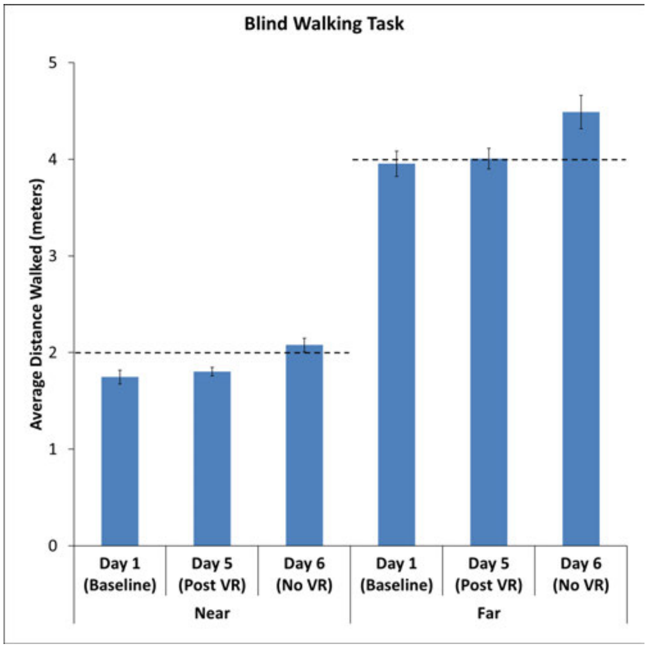

Figure 4: Average distances walked in the blind walking task on days 1, 5, and 6 of the study, for the first trial on each day. No VR immersion occurred on day 6. The dashed horizontal lines represent the distance of the targets from the participants in the near and far conditions, respectively. Error bars represent 95% confidence intervals.

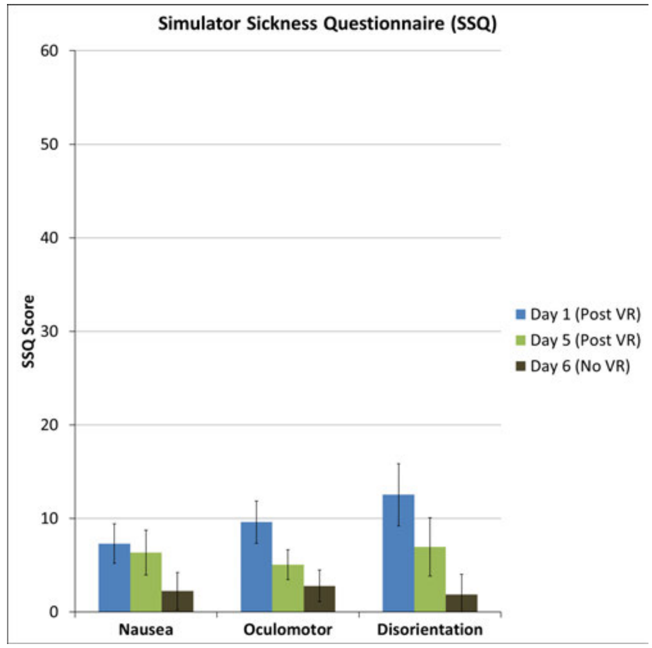

Figure 5: Average SSQ scores on days 1, 5, and 6 of the study. No VR immersion occurred on day 6. The range of the vertical axis was chosen to correspond to that used in [36]. Error bars represent 95% confidence intervals.

with every moment that passes after the removal of the perceptually distorted input. For this reason alone, it is exceedingly unlikely that the difference between day 1 and day 6 can be attributed to an effect of VR usage.

Secondly, in a trial-by-trial examination of the data, it was evident that, with less than 24 hours intervening, participants' average walking distances on the first trial on day 6 resumed exactly where they had left off on the last trial on day 5. In contrast, participants' average walking distances on the first trial on day 5 dropped back down to distances that were comparable to those on the first trial on day 1. This is the pattern of results one might expect if, independent of any other manipulation (e.g., VR use), participants built up comfort and familiarity with walking blindly [34], and these were undone by the passage of time (from day 1 to day 5) but sustained across a rapid return to the task (from day 5 to day 6).

In sum, there is little or no evidence from the psychophysical data collected that VR use by 10 to 12 year olds for half an hour daily across five consecutive days causes any distortions in the representation of three-dimensional space or in hand-eye coordination.

3.3.1 Subjective comfort measures

The scores for the three subscales of the SSQ – nausea, oculomotor discomfort, and disorientation – are shown in Figure 5. What is evident from the figure is that the scores on each of the subscales are not only low by comparison to previously published values, which had mean scores that fell in the 20-50 point range [36], but, moreover, they decrease numerically from day 1 (post VR) to day 5 (post VR). If the SSQ scores on day 6 (24 h post) are taken as a sort of baseline (even though they technically occurred following VR use), then, by comparison, the SSQ scores appear to have been slightly elevated following VR use on days 1 and 5 (all ps $< .05$, with the exception of the SSQ-O day 5 vs. day 6 comparison), but not to a degree that would suggest participants were nauseated, experienced oculomotor discomfort, or were disoriented. Participants also did not exhibit any overt signs of experiencing "simulator sickness" or other forms of subjective discomfort.

In sum, the findings from the SSQ are in line with those from the other measures collected in the present study. The adverse effect probably most commonly associated with VR use [37] proved to be absent, or at least its prevalence was low, in the present study. It is possible that this result may be attributable to the specific collection of content employed, none of which induced excessive amounts of illusory self-motion or required participants to walk around. At the same time, this content was entirely appropriate to the use case evaluated in the present study, to the extent that immersive educational content can be expected to contain varying levels of passive and interactive elements, as well as a "gamification" of learning [6, 28].

4 Discussion

The present study examined short-term use of VR by children ages 10 through 12 years in an educational use case scenario, which was operationalized as 30 minutes of daily VR immersion for five consecutive days. No doubt, this extent of immersion likely exceeds the use one might encounter in a typical educational use case. However, the massed use in the present study ensured that the parameters of use would present a form of stress test for the bodily systems involved: namely the visual system and the vestibular system.

A battery of measures designed to detect the presence of adverse effects on vision, balance and vestibular functioning, spatial representation, hand-eye coordination, and subjective comfort all produced non-anomalous results, suggesting that health or safety consequences are unlikely to occur from moderate use of VR by children in an educational setting.

The present study is the first, to our knowledge, to subject participants to such a comprehensive battery of tests spanning different bodily functions (vision, balance, etc.). Other studies have either assessed fewer dimensions or used smaller sample sizes, or both. Furthermore, the measures used in the current study were adopted from peer-reviewed scientific studies [13, 18–22], and these measures were selected specifically for their relevance to alleged effects of VR usage as described in the popular press [1, 2] and in prior VR research [7–11]. For this reason, the current study's battery of measures, or some combination thereof, could prove useful for future studies that aim to explore potential adverse effects of VR usage. One exception, however, may be the blind walking task, which may be susceptible to the aforementioned task repetition artifact [34]. Future studies, especially those with study designs that entail multiple measurement time points, could consider an alternative measure of far-distance spatial perception, such as a blind tossing task, which past research has found to be comparable to the blind walking task [38].

When considering the current findings in the context of the extant body of literature, it is important to note that past studies finding adverse effects of VR employed technologies that must be regarded as extremely outdated by today's standards [7,10,39,40]. Therefore, it is plausible that the past studies that used these older technologies produced results that are, similarly, outdated. It is also important to note that in some of the past studies that have demonstrated adverse effects with the use of VR, the researchers specifically manipulated the VR content in some fashion to explore the effects of certain types of sensory conflicts on human perceptual or psychomotor systems [5, 8, 9]. In contrast to these past studies, the products tested in the present study represented the state-of-the-art of current commercially available VR products for consumer use, and they were used in their default off-the-shelf condition (i.e., not experimentally manipulated in any way). Their aggregation into a single study affords a generalization of the study results to the present landscape of consumer VR products, when used within a use case scenario that is similar to that of the present study. Moreover, the results of our study are in line with those of other studies – few as they are – of the potential adverse effects of short-term VR use on children [41, 42]. These other studies, which used some of the same VR devices as the current study (including the HTC Vive, Oculus Rift, and Samsung Gear VR), likewise failed to find any substantial effects on the health and well-being of children.

5 Limitations

The current study was designed to assess the potential health and safety ramifications of VR use by users in a particular age range and within a specific use case scenario – namely, short-term VR use by children ages 10 to 12 years in an educational use case scenario. Therefore, potential conclusions that can be drawn from the current research findings must be considered within the context of those study parameters, and those parameters impose natural limitations to the extent to which one can extrapolate the current findings beyond the design of the present study. In future studies, researchers should consider expanding upon these study parameters, such as by testing other age ranges, usage durations, or use case scenarios.

Furthermore, although the present study lacked a control group (i.e., a non-VR group), we consider this study's pattern of results to be readily interpretable without a control group, as the study's within-subject design demonstrated that participants' post-VR performance did not differ statistically from their baseline (pre-VR) performance. The only exception to this was in the blind walking task, in which there was a statistically significant result on day 6 compared to baseline; however, as previously discussed, this result was more likely attributable to a task repetition artifact [34], as opposed to an effect of VR usage. For reasons such as this, in future studies, researchers should consider including a control group, so that any potential effects observed in a VR group can be compared to and assessed within the context of a non-VR group.

Finally, the current study did not address questions related to potential adverse effects of longer-term VR use by children, nor was it designed to address questions related to the potential psychological effects that different types of VR content may have on children. These topics are of widespread potential interest and should be considered as avenues for future investigations.

6 Conclusion

Overall, the findings in the current study indicate that, when used responsibly (e.g., within a safe environment and under adult supervision), modern VR devices can be used safely by children in an educational use case scenario. Indeed, VR has the potential to be a powerful tool, not only within education, but in a wide range of scenarios, each of which may present a unique set of health- and safety-related concerns. Such concerns can and should be examined by future studies, using methodologies that are custom-tailored to the parameters (e.g., demographic, environmental, etc.) that are unique to the use case scenario in question.

References

[1] http://resourcemagonline.com/2018/03/the-negative-side-effects-of-virtual-reality/87052. Accessed: 11/20/2019.

[2] https://www.npr.org/sections/alltechconsidered/2014/08/05/338015854/virtual-realitys-next-hurdle-overcoming-sim-sickness. Accessed: 11/20/2019.

[3] J. Negen, E. Heywood-Everett, H. E. Roome, and M. Nardini. Development of allocentric spatial recall from new viewpoints in virtual reality. *Dev. Sci.*, 21:e12496, January 2018.

[4] S. H. Creem-Regehr, Gill D. M., G. D. Pointon, B. Bodenheimer, and J. K. Stefanucci. Mind the gap: Gap affordance judgments of children, teens, and adults in an immersive virtual environment. *Front. Rob. AI.*, 6:96–110, October 2019.

[5] H. Adams, G. Narasimham, J. Rieser, S. Creem-Regehr, J. Stefanucci, and B. Bodenheimer. Locomotive recalibration and prism adaptation of children and teens in immersive virtual environments. *IEEE Trans. Vis. Com. Graph.*, 24:1408–1417, January 2018.

[6] L. Freina and M. Ott. A literature review on immersive virtual reality in education: State of the art and perspectives. *11th Int. Sci. Conf. eLearn. Soft. Educ.*, 1:133–141, April 2015.

[7] P. Kozulin, S. L. Ames, and N. A. McBrien. Effects of head-mounted display on the oculomotor system of children. *Optometry Vis. Sci.*, 86:845–856, July 2009.

[8] H. Akiduki, S. Nishiike, H. Watanabe, K. Matsuoka, T. Kubo, and N. Takeda. Visual-vestibular conflict induced by virtual reality in humans. *Neurosci. Letters*, 340:197–200, April 2003.

[9] F. A. Biocca and J. P. Rolland. Virtual eyes can rearrange your body: Adaptation to visual displacement in see-through, head-mounted displays. *Presence*, 7:262–277, June 1998.

[10] P. A. Howarth and M. Finch. The nauseogenicity of two methods of navigating within a virtual environment. *Applied Ergonom.*, 30:39–45, February 1999.

[11] K. M. Stanney, R. S. Kennedy, and J. M. Drexler. Cybersickness is not simulator sickness. In *Proc. Human Factors Ergo Soc.*, pages 1138–1142, 1997.

[12] J. W. Miller and J. E. Goodson. Motion sickness in a helicopter simulator. *Aerospace Med.*, 31:204–212, March 1960.

[13] R. S. Kennedy, Berbaum K. S. Lane, N. E., and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *Int. J. Aviation Psych.*, 3:203–220, July 1993.

[14] https://www.wired.com/2015/08/obscure-neuroscience-problem-thats-plaguing-vr/. Accessed: 9/10/2019.

[15] A. Paulus, M. M. Straube and T. Eggert. Vergence-accommodation conflict in virtual reality displays induces phoria adaptation. *J. Neurol.*, 264:16–7, October 2017.

[16] S. V. G. Cobb. Measurement of postural stability before and after immersion in a virtual environment. *Applied Ergonom.*, 30:47–57, February 1999.

[17] H. von Helmholtz. *Treatise on physiological optics*, volume 3. New York: Dover, 1962.

[18] R. Loerzel, L. Tran, and D. A. Goss. Effect of lens power on binocular lens flippers accommodative facility rates. *J. Behav. Optometry*, 14:7–9, January 2003.

[19] S. L. Whitney and D. M. Wrisley. The influence of footwear on timed balance scores of the modified clinical test of sensory interaction and balance. *Ach. Phys. Med. Rehabil.*, 85:439–443, March 2004.

[20] G. M. Halmagyi and I. S. Curthoys. A clinical sign of canal paresis. *Arch. Neurol.*, 45:737–739, July 1988.

[21] J. Long and C. Siu. Two practical tasks for assessing depth perception at near distance – comparison of performance under monocular and binocular conditions. In *Proc. 39th Ann. Conf. Ergo. Soc. Austria*, pages 116–121, 2003.

[22] J. M. Loomis, J. A. Da Silva, N. Fujita, and S. S. Fukusima. Visual space perception and visually directed action. *J. Expt. Psych: Human Percept. Perf.*, 18:906–921, November 1992.

[23] `https://www.lenovo.com/us/en/virtual-reality-and-smart-devices/virtual-and-augmented-reality/lenovo-mirage-solo/Mirage-Solo/p/ZZIRZRHVR01`. Accessed: 11/20/2019.

[24] `https://www.vive.com/us/product/vive-virtual-reality-system/`. Accessed: 11/20/2019.

[25] `https://www.oculus.com/rift/`. Accessed: 11/20/2019.

[26] `https://www.playstation.com/en-gb/explore/playstation-vr/tech-specs/`. Accessed: 11/20/2019.

[27] `https://www.samsung.com/global/galaxy/gear-vr/specs/`. Accessed: 11/20/2019.

[28] A. I. Abdul Jabbar and P. Felicia. Gameplay engagement and learning in game-based learning: A systematic review. *Rev. Educ. Res*, 85:770–779, December 2015.

[29] `https://www.exponent.com/purc/`. Accessed: 11/20/2019.

[30] B.B. Rainey, T. L. Schroeder, D. A. Goss, and T. P. Grosvenor. Reliability of and comparisons among three variations of the alternating cover test. *Ophthal. Physiol. Optics*, 18:430–437, September 1998.

[31] `http://www.stereooptical.com/products/stereotests-color-tests/butterfly/`. Accessed: 11/20/2019.

[32] S. D. Young, B. D. Adelstein, and S. R. Ellis. Demand characteristics in assessing motion sickness in a virtual environment: Or does taking a motion sickness questionnaire make you sick? *IEEE Trans. Vis. Comp. Graph*, 13:422–428, September 2007.

[33] D. Cousineau. Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutor. Quant. Methods Psychol.*, 1:42–45, March 2005.

[34] J. W. Philbeck, A. J. Woods, and J. Arthur. Progressive locomotor recalibration during blind walking. *Percep. Psychophys.*, 70:1459–1470, November 2008.

[35] J. Mollon. After-effects and the brain. *New Sci.*, 61:479–482, February 1974.

[36] R. S. Kennedy, J. M. Drexler, D. E. Compton, K. M. Stanney, D. S. Lanham, and D. L. Harm. Configural scoring of simulator sickness, cybersickness and space adaptation syndrome: Similarities and differences. *Virt. Adaptive Environm.: Applic. Implic. Hum. Perf. Issues*, 1:247, June 2003.

[37] J. J. LaViola Jr. A discussion of cybersickness in virtual environments. *ACM Sigchi Bull*, 32:47–56, January 2000.

[38] C. C. Sahm, S. H. Creem-Regehr, W. B. Thompson, and P. Willemsen. Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Trans. App. Percep.*, 2:35–45, January 2005.

[39] K. M. Stanney and R. S. Kennedy. Aftereffects from virtual environment exposure: How long do they last? *Hum. Fac. Erg. Soc. P.*, 42:1476–1480, October 1998.

[40] S. V. G. Cobb, S. Nichols, A. Ramsey, and J. R. Wilson. Virtual reality-induced symptoms and effects (VRISE). *Presence-Teleop. Virt.*, 8:169–186, April 1999.

[41] D. Yamada-Rice, F. Mushtaq, A. Woodgate, D. Bosmans, I. Douthwaite, W. Harris, R. Holt, D. Kleeman, J. Marsh, E. Midov, M. Mon-Williams, B. Parry, A. Riddler, A. Robinson, D. Rodrigues, S. Thompson, and S. Whitley. Children and virtual reality: Emerging possibilities and challenges. Leeds, UK: University of Leeds. 2017. Accessed at: http://digilitey.eu.

[42] S. G. Ha, K. H. Na, I. J. Kweon, Y. W. Suh, and S. H. Kim. Effects of head-mounted display on the oculomotor system and refractive error in normal adolescents. *J. Pediat. Ophth. Strab.*, 53:238–245, July 2016.