Demographic and Behavioral Correlates of Cybersickness: A Large Lab-in-the-Field Study of 837 Participants

Tiffany Luong*  Adéla Plechata†  Max Möbus*  Michael Atchapero†  Robert Böhm†‡§
Guido Makransky†  Christian Holz*  

† Department of Computer Science, ETH Zürich, Zürich, Switzerland
‡ Department of Psychology, University of Copenhagen, Copenhagen, Denmark
§ Department of Occupational, Economic, and Social Psychology, University of Vienna, Vienna, Austria
‡ Copenhagen Center for Social Data Science (SODAS), University of Copenhagen, Copenhagen, Denmark

Abstract
Cybersickness has been one of the main impediments to the widespread adoption of Virtual Reality for decades. It has been argued that several factors can influence the occurrence of cybersickness, such as technical factors, interaction design, but also users' demographics and their perceived presence. Yet, previous studies had comparably small sample sizes and demographically homogeneous samples; comparisons across studies (e.g., regarding demographic factors) are challenging due to the large variation in the studied virtual environments. In this paper, we address these limitations and report the results of a lab-in-the-field experiment on cybersickness with a large and heterogeneous sample of \( N = 837 \) participants who navigated and interacted inside a virtual environment (ages 18–80, \( M = 29.34, SD = 9.50 \), 431 males, 400 females, 6 non-binaries and other). We found that female participants and participants with lower VR experience were more susceptible to experiencing higher levels of cybersickness. Participants’ cybersickness levels increased with the time spent in VR and with the distance traversed in the virtual world up to a point, above which reported levels declined. We also found a link between higher levels of cybersickness and reduced head motion, as well as between lower levels of cybersickness and more head motion, which led them to explore more of the virtual environment. In contrast to past studies, we did not find any evidence suggesting an effect of age on cybersickness, nor a negative correlation between presence and cybersickness. Based on our results, we derived a model that achieves a mean classification accuracy of 67.1% for two levels of cybersickness using demographic, user experience, and behavioral data in VR.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality

1 Introduction
Virtual Reality (VR) technologies immerse users in computer-generated 3D worlds, allowing full interaction with the displayed environment, and evoking a sense of presence in the virtual space [63]. This sense of presence opens up the high potential for improvement and development in a wide range of application fields, from training [21] and therapy [5, 62] to social networking [51] and entertainment [2], and can even lead to behavioral changes [46, 73].

One of the main challenges holding VR back from wider adoption is “cybersickness,” the bodily discomfort associated with exposure to immersive content [75]. Cybersickness can produce symptoms of nausea, dizziness, and headaches [57]. These symptoms mimic those of motion sickness, but cybersickness and motion sickness are considered distinct due to their difference in actual physical motion [57] and in symptom severity [66, 67, 69]. Cybersickness has been a known condition in users of virtual and augmented reality systems for decades and even longer in simulators [57]. It impacts user experience and can break users’ immersion in the virtual environments, but can also lead to injury or decreased capacity [57].

While influences of technological and interaction factors on cybersickness have been extensively studied (see survey articles for detailed analyses [12, 57, 71]), previous studies have placed little focus on demographic factors [41, 58]. The lack of evidence regarding a quantified impact of characteristics such as age, presence in VR, and participants’ behavior on cybersickness in combination with the typically small and demographically homogeneous sample sizes of participants in laboratory studies [75] constitutes a major limitation of a more comprehensive understanding of cybersickness. Indeed, Peck et al. [50] and Himmelsbach et al. [26] summarized that past human-computer interaction studies have mostly involved male, white, educated, industrialized, rich, and democratic participants, which significantly under-represents the broader population.

In this paper, we report the results of a large-scale VR intervention study in a museum with a demographically more heterogeneous sample of \( N = 836 \) participants (ages 18–30, \( M = 29.34, SD = 9.51 \)), non-males = 405). Participants were immersed inside a virtual environment that simulated a large indoor event inside a ballroom...
during a wedding ceremony as shown in Figure 1. Participants freely navigated the environment and were instructed to complete a series of checkpoints, which led them through multiple rooms, passing roaming people, and finding specific persons to engage in simulated conversations. After completing the VR experience, participants reported their levels of cybersickness and presence in a questionnaire and commented on their VR experience.

Based on the data we collected throughout the study, we derived a cybersickness generalized additive model (GAM) that allowed us to gain insights into occurrences of cybersickness with respect to its demographic and behavioral correlates. Our analysis shows that female participants experienced significantly more cybersickness than male participants and that cybersickness incidence decreased with VR experience. We also found an interaction effect between gender and body ownership. High cybersickness levels for female participants were associated with low body ownership, and high cybersickness levels for male participants were associated with high body ownership. Cybersickness increased with the total duration of experiencing VR, and we found a non-linear relation between the amount of navigation and cybersickness (i.e., total distance traversed in the virtual environment). Overall, cybersickness increased with the amount of navigation in the virtual environment (VE); past a certain point, however, we found a higher amount of navigation to be associated with lower cybersickness scores. Our findings also suggest that participants who reported higher cybersickness ratings moved their heads less while inside VR. Contrarily to past studies, we did not find an influence of age, nor a significant association between presence and cybersickness. Using the variables that had a significant relation with cybersickness, our GAMs achieved a mean accuracy of 67.1% to classify two levels of cybersickness following a user-independent evaluation approach. We conclude with a discussion of our results and put them into perspective.

2 RELATED WORK

2.1 Demographic Correlates of Cybersickness

Past works that studied the influence of demographic information on cybersickness are summarized in Table 1. Studying the influence of participants' demographic information on cybersickness is not novel, but reports regarding the influence of factors tend to contradict [71]. For example, almost as many studies found sex or gender to have a significant effect on cybersickness [22, 28, 29, 47, 65, 69] as studies that did not find any significant effect of these factors [9, 11, 23, 35, 37, 47, 77] (see Table 1).

Overall, female participants were found to experience more cybersickness than males [22, 28, 29, 47, 65, 69, 71]. This gender effect is in line with studies outside VR as female participants were found to experience more motion sickness and visually induced motion sickness than males [19, 65]. Various explanations have been suggested to explain gender differences in cybersickness. Some explained this difference with the social desirability response bias; male participants can under-report illness or belittle their symptoms to better fit society’s gender role expectation [22]. Others attributed the differences to gender dysmorphisms in a seated posture and postural sway [47] or to the non-fitting of the interpupillary distance in VR head-mounted displays (HMDs) [65].

Age was also found to influence cybersickness in past studies [1, 22]. This factor was less studied in past work as recruiting a large sample with senior participants can be more difficult [56]. Overall, past studies reported that young users tend to experience less cybersickness than older ones [1, 22, 35]. These results also corroborate past simulator works in non-immersive settings [49].

Past work agrees that habituation can influence cybersickness symptoms. Howarth and Hodder [27] demonstrated that ten exposures to HMD sessions allow most people to adjust to a system. This habituation phenomenon can also extend to video game experience [56, 76] where lower game experience is usually associated with increased cybersickness incidence [29, 35, 71, 74, 76] (for null effects, see Hakkinen et al. [22]).

There is also a general consensus regarding the influence of past motion sickness history on the incidence of cybersickness symptoms [56, 65, 69]. Motion sickness history is traditionally assessed using the motion history questionnaire (MHQ) [30] or Golding’s motion sickness susceptibility questionnaire [16]. In particular, Stanley et al. [69] found a correlation between cybersickness and the carnival ride motion sickness score item from the MHQ.

In general, past studies tend to hold little statistical power due to the relatively small sample size. There are some notable exceptions, such as Stanney et al. [69]’s study in 2003 that gathered 960 participants in a laboratory setting to study the influence of scene complexity, navigation control, gender, body mass index, and motion sickness history on cybersickness. They found that female participants experienced more severe cybersickness symptoms than males, that body mass index was negatively related to oculomotor symptoms, and that cybersickness symptoms ratings increased with the MHQ susceptibility scores.

2.2 Other Correlates of Cybersickness

Kolasinski [36] published a list of over two-dozen factors in relation to illness exhibited in simulators. In this section, we report findings about presence and other objective metrics. We refer readers interested in hardware or interaction attributes to prior surveys and reviews on this topic [12, 36, 57].

Duration of exposure was overall found to be positively related to cybersickness (episodes and symptom severity) [28, 36, 45, 69], with works suggesting limiting the duration of VR exposure [12, 69]. However, the relation between presence and cybersickness is far more unclear. Some work found that presence and cybersickness were positively associated [34, 39], negatively associated [1, 76, 78], and some did not find evidence that they were associated [9, 72]. In a recent review, Weech et al. [75] conclude that past findings balance in favor of a negative relationship between presence and cybersickness. However, they mention that high-powered studies could help reveal the nature of this relationship.

Regarding objective metrics, past work mostly focused on physiological modalities such as cardiac activity [3, 34, 42, 44, 82], epidermal activity [42, 44], brain activity [8, 34, 71], respiratory activity [13] or eye-activity [6, 34]. Postural stability has also been extensively explored in the literature as a cause [9, 22, 47] or consequence of cybersickness [23, 69, 82]. However, these metrics can only be assessed using external hardware or modules embedded into VR HMDs [40], which are not always accessible to the general consumer because of their cost and setup requirements.

Regarding modalities that can easily be assessed with current VR HMDs, behavioral measures such as the amount of head motions or interaction in the VE, which is related to interaction and engagement factors [70], have so far received much less attention as potential correlates of cybersickness. We aim to close this gap.

3 METHODS

3.1 Context of the Study

We conducted the experiment in the National Museum of Berlin over 16 days, from November 11 to 18, 2021. Displays in the museum advertised the study during that time for recruiting purposes. We also advertised and recruited participants on social media networks one month prior to the start of the experiment. In total, 909 participants volunteered to take part in the study or were actively approached in person as they visited the museum.

We designed the experiment for several independent objectives. The first was an entertaining and informative experience about COVID-19 spread and options for vaccination. The details, methods, and results can be found in a submission to a journal in psychology [52]. The second objective of the experiment, investigated in
For full details on the implementation of the experiment, including the small variations between the three environments, we refer the reader to our previous work [52].

3.2 Virtual Environments

In the VR application, participants were guided through different virtual environments where they had to perform simple interactions. We designed three different versions of the VR experience. The versions participants experienced differed by the presence of gamified elements and/or empathy content. The different versions did not influence the final cybersickness model, however. Each participant experienced one version of the VR experience.

In all versions of the virtual environment, participants embodied an elderly character (gender-matched) who could not get vaccinated against COVID-19 due to medical reasons and was invited to the wedding of his/her granddaughter. We chose elderly avatars as we also aimed at communicating the impact of vaccination on COVID-19 spread for at-risk individuals [52]. Participants were first exposed to an apartment scene in which they were instructed to wash their hands in front of a mirror where they could see their avatar. Then, they received a wedding invitation from their granddaughter.

In the next scene, participants were free to navigate in the 20 × 20 m virtual ballroom (trajectory was not predefined) as shown in Fig. 1. They were guided through different objectives with narration, including signing the guest book, placing a present on a table, and talking to the bride. Participants completed all objectives while following the instruction to actively avoid the other guests, for whom the trajectories were predetermined. We piloted these paths to ensure they were sufficiently complex not to be predicted by participants. When participants accomplished an objective (by reaching a landmark on the floor or by colliding their virtual hand with a highlighted object), the next objective was indicated through a large text box, arrows, and highlighted landmark in the virtual environment.

For full details on the implementation of the experiment, including small variations between the three environments, we refer the reader to our previous work [52].

Table 1: Summary of our study and past VR studies’ findings on the influence of demographic factors on cybersickness. The sources are ordered from most recent (2022) to most ancient (1996). NA=unknown, y=yes, n=no, (empty cells)=not assessed. MS hist.: motion sickness history. SSQ: simulator sickness questionnaire [31]. FMS: fast motion sickness scale [32]. Cust.:custom scale.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>N</th>
<th>Age Range</th>
<th>M</th>
<th>SD</th>
<th>Non-males ratio</th>
<th>VR hardware</th>
<th>Interaction</th>
<th>Context</th>
<th>Seated</th>
<th>Duration (min)</th>
<th>Cyber sickness measure</th>
<th>Influence of</th>
<th>Gender</th>
<th>Age</th>
<th>Game exp.</th>
<th>MS hist.</th>
</tr>
</thead>
<tbody>
<tr>
<td>[47]</td>
<td>2017</td>
<td>36</td>
<td>NA 20.72 0.85</td>
<td>50%</td>
<td>HMD</td>
<td>joystick</td>
<td>lab</td>
<td>n</td>
<td>NA</td>
<td>15–SSQ</td>
<td>y</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[48]</td>
<td>2017</td>
<td>36</td>
<td>NA 22.72 3.56</td>
<td>50%</td>
<td>HMD</td>
<td>joystick</td>
<td>lab</td>
<td>NA</td>
<td>NA</td>
<td>SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[56]</td>
<td>2014</td>
<td>20</td>
<td>NA 18–31</td>
<td>15%</td>
<td>HMD</td>
<td>joystick</td>
<td>lab</td>
<td>n</td>
<td>NA</td>
<td>SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[23]</td>
<td>2007</td>
<td>41</td>
<td>NA 49%</td>
<td>HMD</td>
<td>NA</td>
<td>lab</td>
<td>NA</td>
<td>30</td>
<td>60 NA</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[35]</td>
<td>2006</td>
<td>387</td>
<td>9–60</td>
<td>NA</td>
<td>NA</td>
<td>CAVE</td>
<td>NA</td>
<td>NA</td>
<td>SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[22]</td>
<td>2003</td>
<td>60</td>
<td>18–41</td>
<td>26.7</td>
<td>HMD</td>
<td>mouse</td>
<td>lab</td>
<td>y</td>
<td>15–60 SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[28]</td>
<td>2001</td>
<td>60</td>
<td>18–40</td>
<td>30%</td>
<td>HMD</td>
<td>mouse, treadmill</td>
<td>lab</td>
<td>n</td>
<td>15–21+ SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[68]</td>
<td>1999</td>
<td>34</td>
<td>NA 25.79 4.72</td>
<td>41%</td>
<td>HMD</td>
<td>mouse</td>
<td>lab</td>
<td>y</td>
<td>30 SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[37]</td>
<td>1998</td>
<td>40</td>
<td>19–46</td>
<td>22.7</td>
<td>HMD</td>
<td>mouse</td>
<td>lab</td>
<td>y</td>
<td>20 SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[36]</td>
<td>1996</td>
<td>40</td>
<td>19–46</td>
<td>22.7</td>
<td>HMD</td>
<td>mouse</td>
<td>lab</td>
<td>y</td>
<td>20 SSQ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3 Apparatus

The support application was developed using Unity and was presented using Oculus Quest 2 in a standalone setting (72 FPS on average). Participants interacted and navigated in the virtual environment using both Oculus Touch. They could move forward and backward using either the controller’s joystick in the direction of their headset (i.e., steering locomotion). When moving, the speed was set to be constant at .75 m/s (not dependent on the joystick motion amplitude) and piloted before the user study with nine persons several times to adjust the game difficulty. There was no deceleration or acceleration in the motion. Participants could pick and drop objects by colliding their avatars’ hands with the 3D models.

The avatars and newlywed models were designed by a professional artist, and the wedding guests were RocketBox avatars [18]. We used voice recordings in either English or German by professional voice actors/actresses for all instructions and interactions with virtual avatars.

3.4 Measures

We considered the following data in the analyses:

Cybersickness: We asked participants to report their level of motion sickness after the study using the following question: “How motion sick (nauseous or dizzy) did you feel during the wedding experience?” (Scale from 1–Not at all, to 5–Extremely).

Demographic: Before the study, we assessed the participants’ gender (1–male, 2–female, 3–non-binary, 4–other), age, VR experience (“How many times have you used Virtual Reality before?” 1–“Never”, 2–“1–3 times”, 3–“4–10 times”, 4–“11–20 times”, 5–“More than 20 times”), and language preference in the VR application (English, German).

User experience (UX): After the study, participants reported their level of presence and social presence (adapted from [43]): “While I was at the wedding, I had a sense of “being there”, “I had a sense that I was interacting with other people at the wedding rather than a computer simulation”, scales from 1 (strongly disagree) to 5 (strongly agree), body ownership (from [17]: “I felt as if the virtual body I saw when I looked down was my body”, scale from 1–strongly disagree, to 7–strongly agree).
Behavioral: We assessed the total time spent in VR, the accumulated HMD movement (based on the accumulated difference in HMD position in the tracking space), the accumulated HMD angular movement (based on the accumulated difference in HMD rotation in the tracking space), and the accumulated distance traversed in the VR environment.

Table 2 shows the key sample characteristics.

Table 2: Studied variables and characteristics of the analyzed study sample. Continuous variables are summarized as $M$ (SD).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data ($N = 837$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age [18–80] (in years)</td>
<td>29.34 (9.50)</td>
</tr>
<tr>
<td>% Male</td>
<td>51.5% ($n = 431$)</td>
</tr>
<tr>
<td>% Female</td>
<td>47.8% ($n = 400$)</td>
</tr>
<tr>
<td>% Non-binary and other</td>
<td>.7% ($n = 6$)</td>
</tr>
<tr>
<td>Prior VR experience [1–5] (median)</td>
<td>1–Never</td>
</tr>
<tr>
<td>% German speakers</td>
<td>60.1% ($n = 503$)</td>
</tr>
<tr>
<td>Cybersickness [1–5] (median)</td>
<td>2</td>
</tr>
<tr>
<td>Presence [1–5] (median)</td>
<td>4</td>
</tr>
<tr>
<td>Social presence [1–5] (median)</td>
<td>3</td>
</tr>
<tr>
<td>Body ownership [1–7] (median)</td>
<td>5</td>
</tr>
<tr>
<td>Time spent in VR (in s)</td>
<td>701.98 (215.83)</td>
</tr>
<tr>
<td>HMD movement (in m)</td>
<td>31.32 (13.13)</td>
</tr>
<tr>
<td>HMD angular movement (in °)</td>
<td>9905.10 (4434.16)</td>
</tr>
<tr>
<td>Distance in VR (in m)</td>
<td>156.53 (48.01)</td>
</tr>
</tbody>
</table>

3.5 Participants and Experimental Procedure

In total, 909 participants took part in the experience. 54 participants were excluded from our final sample because they dropped out of the experience due to factors such as time constraints, incomplete or invalid data, or discomfort.

Of the resulting 855 participants, we excluded 18 additional participants. We excluded nine participants who accumulated an HMD angular movement of more than 25000°, since these observations were larger than the 3rd sample quartile plus 1.5 times the inter-quartile range and were found to have a large influence on the model fit (see supplementary materials). For similar reasons, we excluded nine further participants who spent more than 2250 s in VR as part of the experiment (see supplementary materials). Furthermore, we analyzed the data from the participant who reported “other” for gender with the “non-binary” group.

The final sample consisted of $N = 837$ participants, whose ages ranged from 18 to 80 years (Table 2 lists the sample characteristics).

Participants could choose to complete the simulation and questionnaires either in English or in German. They first filled out an informed consent form and a pre-treatment questionnaire. Then, they were randomly assigned to one of the three VR conditions and spent on average 11.7 min ($SD = 3.6$ min) in the VR simulation. After the VR experience, participants filled out a post-treatment questionnaire. They were compensated with a small gift (a cup worth 4.50 EUR) and an additional 5 EUR voucher for the museum gift shop.

The study was pre-registered on November 9, 2021, prior to the data collection. It was approved by the relevant Institutional Review Board, approval number IP-IRB/02092021.

4 Results

4.1 Data Analysis Methods

In this paper, we analyze which demographic, UX and behavioral factors are significantly related to cybersickness. First, we analyzed the relations between variables using Spearman correlations. Then, we modeled cybersickness using a proportional odds logistic regression (POLR) model, since it was reported as an ordered factor on a scale from 1 to 5. In particular, we modeled cybersickness using a POLR GAM, to account for potentially non-linear effects.

Due to their flexibility, GAMs have been a popular modeling technique for complex, high-dimensional data over the last decades and have been constantly enhanced in software such as R [24, 53, 61, 79, 80]. GAMs are particularly frequently used in the fields of environmental, geological, and ecological science [14, 20, 54].

We constructed the POLR GAM for cybersickness using the “MGCV” library [79] in R [53]. The general model included all of the demographic, behavioral, and UX variables indicated in Section 3.4, as well as the different application versions (see Section 3.2). Variables were selected using a sequential backward elimination. The effects with the largest p-values were removed from the model until all effects were statistically significant at a confidence level of .10. Initially, we fitted all effects of not-categorical variables as non-linear effects. If an effect then showed no evidence of non-linear behavior, e.g., was fitted close to a straight line, we subsequently fitted it as a linear effect. Non-linear effects were fitted as thin-plate regression splines, as generally recommended [79, 81]. We included a further penalty on the null space of the regression splines shrinking effects to zero and effectively removing them from the model in case of low explanatory power. This approach allows for variable selection, reduces overfitting of the non-linear effects, and, thus, generally promotes a more conservative fit. In addition to non-linear effects, we also tested for second-order interaction effects. We fitted the GAM using marginal likelihood as the smoothing selection criterion since this results in p-values with the best behavior compared to alternative smoothing selection criteria according to the authors of the used software package “MGCV” [79]. As proposed, we verified that no pair of regression splines exceeded a concavity of .50 [25].

Finally, we evaluated the cybersickness classification performances using the variables retained in the final model and leave-2-participants-out cross-validation to further test the robustness of the relation between the selected variables and cybersickness.

4.2 Correlation Results

The reported cybersickness score correlated statistically significant with participants’ gender ($\rho = 0.28, p < .001$), previous VR experience ($\rho = -0.15, p < .001$), total duration in VR as part of the experiment ($\rho = 0.08, p < .05$) and accumulated HMD angular movement ($\rho = -0.08, p < .05$). No other correlations were statistically different to 0, at a confidence level lower than .05. A full correlation matrix between all collected variables can be found in Fig. 2.
4.3 Cybersickness Model

In the following sections, we will discuss the variables that were calculated to have a statistically significant effect on cybersickness in the POLR GAM. Following the procedure indicated in Section 4.1, we obtained the following model:

\[
\text{cybersickness} = g(\eta),
\]

with \( \eta_i = \text{gender}_i + \text{VR}_{\text{experience}_i} + \text{body}_{\text{ownership}_i} + \text{time}_{\text{in VR}} + \text{f}_j(\text{hmd}_{\text{rotation}_i}) + \text{f}_2(\text{distance}_{\text{in VR}}) \) for \( i = 1, \ldots, n, \)

where \( f_j \) are smooth functions representing the effects of \( \text{hmd}_{\text{rotation}} \) and \( \text{distance}_{\text{in VR}} \) for \( j = 1 \) and \( j = 2 \), respectively.

If \( \tau_2, \tau_3, \tau_4 \in \mathbb{R} \) are three distinct points on the real line, \( g \) takes the following form for an ordinal response with 5 levels \( 1, \ldots, 5 \):

\[
g(\eta) = \begin{cases} 
1, & \text{if } \eta < -1 \\
2, & \text{if } -1 \leq \eta < \tau_2 \\
3, & \text{if } \tau_2 \leq \eta < \tau_3 \\
4, & \text{if } \tau_3 \leq \eta < \tau_4 \\
5, & \text{otherwise}
\end{cases}
\]

For our final model (POLR GAM):

\[
\tau_2 = 0.77; \quad \tau_3 = 2.04; \quad \tau_4 = 4.14
\]

The partial dependence plots for each variable can be found in Fig. 3.

4.3.1 Demographic and UX Factors

We found a statistically significant effect of gender on the reported cybersickness score (see Fig. 3). Female participants were calculated to have increased chances of experiencing cybersickness compared to male participants \( (p < .0001, \chi^2 = 24.56) \), matching previous observations of Spearman correlations. While participants self-identifying as non-binary or other were also calculated to experience cybersickness more easily than male participants, this difference was not statistically significant. We found previous experience in VR to significantly decrease the chances of experiencing cybersickness \( (p < .0001, \chi^2 = 23.97) \) again matching previously observed Spearman correlations. For female participants, our findings also showed that decreases in experienced body ownership during the experiment significantly increased the chances of observing high levels of cybersickness \( (p < .01; \text{see Fig. 3}) \). For male participants, however, the trend was reversed as increased body ownership tended to increase the expected cybersickness measure \( (p = .10; \text{see Fig. 3}) \). For participants identifying as non-binary or other, the effect of body ownership was not statistically significant \( (p = .92) \).

We did not find any significant effect of age nor presence on cybersickness (see Fig. 4).

4.3.2 Behavioral Variables

Furthermore, we found multiple statistically significant effects connected to the behavior of participants during the experiment: total time spent in VR, total distance travelled in VR, and accumulated HMD angular movement. Increases in total time spent in VR increased the chances of experiencing cybersickness linearly \( (p = .01, \chi^2 = 6.61) \). Increases in accumulated HMD angular movement decreased the expected cybersickness measure non-linearly \( (p < .01, \chi^2 = 9.34) \) as can be seen in Fig. 3. The fitted effect decreased steepest in between \( \approx 10,000^\circ-15,000^\circ \). Total distance traversed in VR also showed evidence of a non-linear effect \( (p < .01, \chi^2 = 9.43) \). We observed an effect similar to a negative quadratic parabola: traversing only short distances in VR (less than 100 m) was calculated to decrease the expected cybersickness measure, as well as traversing long distances (more than 300 m). We found traversing distances around 200 m to maximize the chances of observing cybersickness.

We did not find that other behavioral variables were significantly related to cybersickness.

4.4 Classification Performances

We explored whether the variables retained in the final GAM (Equation 1) could be used to classify cybersickness by performing a leave-2-participants-out cross-validation (1000 iterations, balanced test sets). We compared the performances of POLR GAMs, support vector machines (SVM), and random forests (RF; 500 estimators and maximum depth of 5). The evaluation was performed for 2 classes of cybersickness levels \((1,2) \) vs. \((3,4,5)\), and the training sets were balanced using Synthetic Minority Oversampling Technique (SMOTE) [7].

Table 3 summarizes our results.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Cybersickness ratings ((1,2)) vs. ((3,4,5))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric</td>
<td>Accuracy</td>
</tr>
<tr>
<td>POLR</td>
<td>67.1%</td>
</tr>
<tr>
<td>SVM</td>
<td>64.5%</td>
</tr>
<tr>
<td>RF</td>
<td>59.3%</td>
</tr>
</tbody>
</table>

The results show that the POLR GAMs overall outperformed the SVM and RF models with a mean accuracy of 67.1%.

5 Discussion

In this paper, we conducted a study and assessed participants’ cybersickness level, as well as their demographic information and behavioral data. The richness of this analysis stems from the exceptionally large \((N = 837)\) and diverse \((SD_{\text{age}} = 9.50, \text{non-male}_{\text{ratio}} = 48.5\%)\) sample of participants who were recruited in the wild in a museum, outside laboratory settings. Below we summarize the results on factors being associated with participants’ self-reported levels of cybersickness; we put them into perspectives of the previous literature.

5.1 Demographic and UX Correlates

Female participants reported significantly higher levels of cybersickness than male participants, which corroborates most of the past findings \([22, 28, 29, 47, 65, 69]\). Stanney et al. [65] explained the gender effect with the non-fit of the IPD and Munafo et al. [47] with participants’ postural sway, which we did not control in this study. The HMD was calibrated vertically using the Oculus Quest 2 procedure, but the distance between the lenses was not adjusted between participants. Another potential explanation is social desirability biases linked to society’s gender role expectation \([10, 15, 22]\). Out of the 431 male participants, none reported the maximum level of cybersickness, i.e., 5. Given the exceptionally large and diverse sample, these findings still strengthen past results that found that female users are more susceptible to experiencing cybersickness than male users, also using the most widespread VR HMD on the market in 2022.

In contrast to past studies \([1, 22, 35]\), we did not observe a relation between age and cybersickness (see Fig. 2 and the supplementary materials). Some of these previous studies were brief reports and did not report important details (e.g., age mean and SD, test statistics) or were limited by low sample size. Taken together, previous estimates are likely to be less accurate, over- or underestimating the true effect.
Our results show that an increase in past VR exposures decreased the level of cybersickness reported. Few studies assessed participants’ VR experience (they mostly measured game experience [29, 35, 76]), and the only work that assessed it did not find a significant effect of VR experience on cybersickness [22]. Our results can mainly be explained by the phenomena of habituation [27, 59]. Additionally, it is more likely that people who experienced cybersickness in the past will less self-select into an experience using VR HMD again.

We did not find a correlation between presence (environment and social) and cybersickness (see Fig. 2 and 3). There have been debates about such a relation in the past [9]. A recent review settled for a negative correlation between presence and cybersickness but the authors highlighted the lack of “high-powered studies” [75]. The null correlation might be due to a complicated causal-inference relation on/of cybersickness for different participants. Weech et al. [75] mentioned that a positive relationship could occur because of the mutual impact of immersiveness—the sensory “submersion” experienced by a user [4] or vection—the illusion of self-motion [60] on presence and on cybersickness. These factors are sometimes considered as preconditions for users to experience high levels of presence [43, 60, 64] or high levels of cybersickness (through sensory conflicts) [38, 55]. In contrast, a negative relation between cybersickness and presence can be explained by a top-down relation, i.e., cybersickness could have played a distracting role in the VR experience, suppressing the attention required to achieve high levels of presence [48, 60, 72, 75].

The results also show an interaction effect between gender and body ownership on cybersickness. Female participants with lower body ownership experienced significantly higher degrees of cybersickness, while male participants with higher body ownership tended to experience higher degrees of cybersickness. A similar interpretation as the one for the positive or negative relation between presence and cybersickness can be given for this interaction effect. On the one hand, the positive association between body ownership and cybersickness for male participants could originate from an external factor such as vection or immersiveness. The illusion of self-motion requires the sensorimotor control to be convinced that the visual motion is in line with one’s own body motion [60]. In that sense, body ownership is even more concerned with vection than the overall sense of presence. On the other hand, cybersickness might have prevented female participants to experience the illusion of body ownership. A past work found that female participants experienced a higher onset of cybersickness symptoms than males [23]. This might also have been the case in our study, which could have impacted females’ user experience at an earlier point. Further work could focus on studying this potential difference in cybersickness and body ownership outcome, based on when cybersickness occurred.

5.2 Behavioral Correlates

We found a positive association between the time spent in VR and cybersickness, corroborating past results [28, 36, 45, 69]. Interestingly, we found a negative non-linear association between accumulated HMD angular movement and cybersickness ratings. This effect could seem counter-intuitive from a bottom-up point of view as one could expect a higher amount of interaction and motion to increase the level of cybersickness. The most plausible explanation is that the relation originates from cybersickness (i.e., top-down effect). Participants who experienced cybersickness might have decreased their head movement in VR to avoid worsening their sickness symptoms while a higher amount of HMD rotation...
translated to a lower chance of cybersickness occurrence and higher confidence in interacting in the VR environment (looking around). There are parallels to draw with past results. Indeed, Steinicke et al. [70] found that looking in the direction of the walk indicated that users felt insecure, whereas a free-look-around revealed that users felt safe. Further studies could focus on these interpretations.

Furthermore, the total distance traversed in VR is also significantly associated with the level of cybersickness. More specifically, the level of cybersickness is positively related to the total distance traveled in VR up to a point (~200 m), above which the reported levels declined. Some interpretations: (1) bottom-up in the first part, the cause of this relationship might originate from the interaction, where the more users perform navigation interaction, the higher is their level of cybersickness; (2) top-down in the second part, non-sick participants were more at-ease with interacting in VR and felt free to explore the VR environment to their heart content, similarly to the HMD angular movement results. These effects require further investigation.

5.3 Limitations and Future Work

Our results revealed interesting insights regarding the incidence of cybersickness in—what could be considered a regular VR application showcase for passerby visitors in a museum. We discovered significant relations between behavioral metrics and cybersickness, which pave the way for further research. In particular, our results suggest that while longer VR exposure augments the risk to experience high levels of cybersickness, a large amount of HMD rotation and exploration in VR translates to low levels of cybersickness. Consequently, an interesting question to investigate would be: does cybersickness decrease users’ head rotation? Furthermore, we only assessed sum metrics for behavioral measures. The joint effect of HMD rotation and users’ navigation was not sampled and could be further investigated to better understand sensory conflicts in VR in the future. Overall, a big question stems from these results: how can we disentangle the complex relation between vection, presence, body ownership, users’ behavior, and cybersickness? Those are still unanswered questions that would require additional experiments.

An important limit in our method is the assessment of cybersickness. We used a non-standardized single state question to assess cybersickness, as opposed to the widely spread simulator sickness questionnaire (SSQ) [31]. We did not use the SSQ as the user study was constrained in time (i.e., < 20 min per participant) as part of an exhibition in a museum. In the question, we specified nausea and dizziness symptoms as those weight the most with general discomfort in the SSQ final score [31]. Still, it is unclear from our assessment to what extent participants felt either or both and some symptoms such as ocularmotor ones were not assessed. A shorter standardized scale for VR cybersickness would benefit future large-scale VR experiments and longer questionnaires such as the SSQ or VRSQ [33] could improve the diagnosticity for cybersickness symptoms.

We reported the demographic, UX, and behavioral factors that were significantly related to cybersickness in our study. Using POLR GAMs and machine learning approaches, we reached a mean classification accuracy of 67.1% for two levels of cybersickness, highlighting the potential of the identified variables to predict cybersickness.

Future work could consider several lines for improvement. Especially, physiological modalities have shown promises to detect cybersickness [8, 13, 34, 44, 71]. In our study, we only considered data that could be obtained from the standalone Oculus Quest 2 out of the box. Future work could attempt to establish a large dataset on cybersickness studies in VR with demographic, behavioral, and physiological data to better understand the individual variability of cybersickness and to detect cybersickness in real-time.

Overall, our results are reassuring as they support past findings with a smaller sample size regarding the effect of gender, VR experience, and time in VR. However, they also contradict past work on the relation between cybersickness and age and presence. Therefore, we argue that large sample sizes should be employed whenever possible in future cybersickness studies, especially when demographics are studied because of the large individual variabilities.

6 Conclusion

In this paper, we conducted a large lab-in-the-field study in a museum with \( N = 837 \) participants and analyzed the relation between cybersickness and demographic, user experience, and behavioral measures.

Our results show that female participants and participants with low VR experience experienced higher levels of cybersickness. High cybersickness level for female participants was associated with low body ownership, and high cybersickness level for male participants was associated with high body ownership. We confirm that a longer time in virtual reality (VR) and an increased amount of navigation in VR—until a certain point, increase the risk of experiencing higher levels of cybersickness. Our results also show a positive association between low cybersickness level and increased amount of head rotation, as well as between low cybersickness level and increased distance traversed in the immersive virtual environment. Unlike past work, we did not find an effect of age on cybersickness, nor a significant association between presence and cybersickness. Using the variables that had a significant relation with cybersickness, our models achieved a mean classification accuracy of 67.1% for two levels of cybersickness on unseen participants.

Taken together, our results suggest that female participants and participants with low VR experience are more susceptible to experiencing cybersickness. Some factors such as the time spent in VR and the quantity of navigation in VR can also increase the risk of cybersickness incidence. Our findings suggest that participants who experienced low levels of cybersickness moved their heads and explored the immersive virtual environment more. Those results pave the way for future experiments aiming at understanding the relation between cybersickness, users’ behaviors, vection, presence, and body ownership to better predict cybersickness occurrences.
ACknowledgments

This research was funded by the European Institute of Innovation and Technology (EIT Health, Grant No. 210836). We would also like to warmly thank the National Museum of Berlin and the students from the University of Copenhagen who helped conducting this study.

References


