2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)

# Field Dependence as a Predictor of Cybersickness Dropout

Brady Li*
Northeastern University
Boston, Massachusetts, USA

Zachary Gaydos†
Virginia Tech
Blacksburg, Virginia, USA

Wallace Lages‡
Northeastern University
Boston, Massachusetts, USA

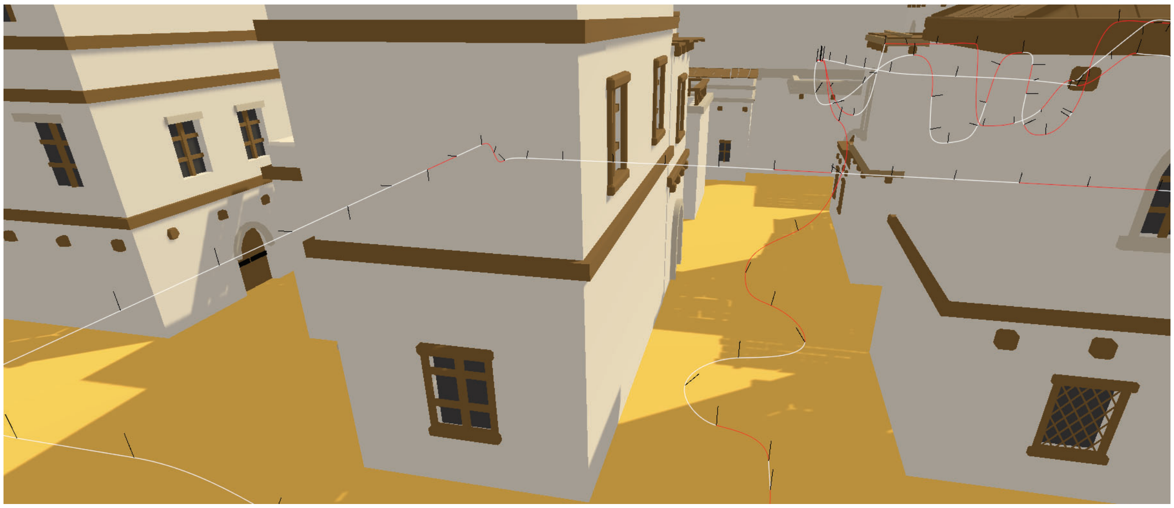

Figure 1: Virtual reality roller coaster used to induce cybersickness. The track outline in white and red was not visible to participants.

**ABSTRACT**

One of the primary barriers to the widespread adoption of virtual reality (VR) is cybersickness, a form of visually-induced motion sickness caused by the mismatch between visual and vestibular cues. Our study investigates the role of field-dependence on cybersickness susceptibility. We used a virtual reality roller coaster simulation to induce sickness and measured the time for participant dropout. Field dependence was measured using the Group Embedded Figure test and two variations of the Rod and Frame test. A survival analysis indicated that field dependence can be an effective measure of individual differences in cybersickness. Field-dependent individuals have a 37% increased risk of dropping out and females showed an increased risk of 126%. We did not find a significant effect from age or sleep quality. The study also confirms that the dropout rates paradigm can effectively identify susceptibility to cybersickness and may be preferable to questionnaires. Psychological tests that measure these variables may eventually blend seamlessly in virtual environments and allow for greater prediction of cybersickness susceptibility.

**Index Terms:** Human-centered computing—Human Computer Interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—HCI design and evaluation methods—User models

## 1 INTRODUCTION

Although virtual reality (VR) head-worn displays have improved, cybersickness continues to be a barrier to entry and regular usage. It affects a large percentage of VR users: between 40% to 70% of users experience symptoms within 15 minutes [35, 51]. Cybersickness symptoms are similar to traditional forms of motion sickness, including general discomfort, nausea, disorientation, vertigo, and vomiting, which can last after the experience [45, 70].

To improve the user experience, cybersickness research has been looking into the causes, developing models for prediction, and creating symptom mitigation techniques [16, 22, 26, 41, 44, 46, 69]. Although progress has been made, the complexity of the condition, along with mixed findings arising from different study designs, still challenge our understanding of cybersickness [26, 27, 52].

Prior research has indicated that individual differences may explain a large part of the variation in susceptibility and intensity of the symptoms. Some of the individual differences identified as potential contributors include sex, age, and personality [27, 61]. Exploring other differences in line with theories of cybersickness may result in more accurate predictions and a better understanding of cybersickness as a whole.

In this paper, we report on the results of an exploratory study investigating the relationship between cybersickness and field dependence, a cognitive style that describes how likely individuals are to use the visual context rather than internal cues as a reference for behavior. Field dependence has been traditionally assessed with the Group Embedded Figure Test (GEFT) [84] or the Rod and Frame test (RFT) [81]. The GEFT asks the participant to find a geometric image embedded in others, while the RFT asks the participant to align a rod vertically against a tilted background. In summary, both tests measure how much the background influences the outcome of the task. Field-dependent individuals are more likely to incorporate the background, while field-independent individuals are more likely to ignore the background. Participants are then ranked according to the number of correct answers to the tasks.

If cybersickness is caused by a mismatch of expected signals from the vestibular organs, field dependence should be a predictor of cybersickness. The reasoning is that a mismatch requires individ-

*e-mail: li.bra@northeastern.edu
†e-mail: zachg@vt.edu
‡e-mail: w.lages@northeastern.edu

2473-0726/25/$31.00 ©2025 IEEE
DOI 10.1109/ISMAR67309.2025.00140

uals to give enough relative weight to the background so that it is perceived as a significant source of conflict. Field dependence can be seen as a reflection of those weights: extremely field-independent people can completely ignore the background (low weight), while extremely field-dependent people cannot prevent considering it (high weight). As a result, one would expect that field dependence will be correlated with the likelihood of becoming sick.

Instead of predicting discomfort scores, we focus on directly predicting how long an individual will stay in the experience. The duration is a useful metric for researchers and practitioners since it allows an estimate of experience time [7]. We formulated and compared prediction models incorporating field dependence along with sex, age, and sleep. Our final model aligns with previous results on sex and extends it to indicate that field-dependent individuals have a 37% higher risk of dropping out of a sickness-inducing virtual reality experience.

This paper's contributions to the understanding of cybersickness include:

1. A comparison of different methods to measure field dependence. We show that the Rod and Frame test in a virtual empty room has higher predictive power than the Group Embedded Figure test or the Rod and Frame in a furnished virtual room.

2. A comparison between a cybersickness questionnaire and a drop-out rate as a measure of cybersickness. We show the reports are not a good predictor for participant drop-out.

3. The evaluation of a time-to-event model that includes field dependence as a predictor for cybersickness. We show that models including sex have better explanatory power than models with age or self-reported sleep.

4. A new metric for obtaining field dependence measurements from the Rod and Frame test. We show that Vingerhoets's harmonic model can be used to build a coherent measure that offers enough explanatory power for predictive models.

This paper is organized as follows: section 2 summarizes the current theories and prior work on cybersickness predictors. Section 3 describes the methodology used in the study, including stimulus and measures. Section 4 describes the results. Section 5 discusses the results. Section 6 discusses limitations and future work.

## 2 Background

Here, we review existing literature on cybersickness, theories, explored methods, contributing factors, and plausible connections to field dependence.

### 2.1 Cybersickness

Cybersickness (or simulator sickness) is considered a form of motion sickness [12] with symptoms and physiological changes similar to sickness in cars, space, and other simulators [45]. Cybersickness symptoms include disorientation (dizziness, vertigo, difficulty focusing), nausea (stomach awareness, increased salivation, nausea), and oculomotor symptoms (eyestrain, headache, blurred vision). Unlike other forms of motion sickness, cybersickness is visually-induced and seems to have a different symptom profile, with disorientation symptoms being more common [64].

There are numerous theories regarding the causes of cybersickness and other visually-induced motion sickness [55]. They include several variants of sensory conflict (e.g., neural mismatch, vection, subjective vertical, and rest frame conflicts), eye muscle proprioception, increased postural instability, and a mismatch between the user's head and virtual pose [13, 55, 62, 77]. However, the most accepted theory is that the symptoms are a result of a neural mismatch between vestibular and visual cues (i.e., the movement seen by the eyes disagreeing with the expected signals from organs related to balance) [13, 21, 22, 30, 42, 46, 62]. In VR, this is typically caused by having a stationary user subject to visual movement cues generated by lag, virtual locomotion, or some type of uncontrolled motion (e.g., riding rollercoaster) [16, 55, 73].

Individual differences such as experience with VR, susceptibility to classical motion sickness, age, visual sensitivity, and sex have been observed to influence cybersickness severity [22, 41, 44, 46, 69]. Understanding the role of individual differences can allow specific mitigation techniques and experiences for different groups. In this section, we will review the individual factors considered in this study. We refer interested readers to the comprehensive review from Mittelstaedt [50] for additional factors.

#### 2.1.1 Sex

Sex has been consistently found to affect cybersickness. Prior research indicates that females are more prone to suffer from both cybersickness and motion sickness [26, 30, 38, 51, 69]. There have been numerous proposals of mediating variables and theories to explain why females seem to get sicker; these include hormonal differences, postural stability, interpupillary distance, gaming experience, presence, and even cultural factors and experimenter bias [30, 39, 40].

Although some studies show no significant difference between sex and cybersickness [27, 46], this has been attributed to weak stimuli or to small sample sizes [30, 39]. Luong et al. [38] found that low sense of presence in women led to higher sickness, whereas high sense of presence in men led to higher sickness. This difference highlights the need to consider the many existing influences when designing and studying new factors.

#### 2.1.2 Age

Prior research has indicated that younger individuals, particularly ages 1-12, are more susceptible to motion sickness, with older adults being less and less susceptible towards 50 years old [37, 63]. This early susceptibility is thought to be a result of a lack of development in the vestibular and visual cues and a lack of overall exposure [23]. However, other studies have shown conflicting results, with some indicating that older adults are more susceptible and experience greater cybersickness [5]. Petri et al. [58] found that female participants over 60 years old and unfamiliar with the content of the experience suffered greater cybersickness. Some studies have also shown no effect with age [27, 38]. Many studies on this factor are from driving simulators [29, 33, 44], which may explain some of the results.

Dilan et al. [18] found that older adults experience a higher sense of presence and lower cybersickness in various VR experiences. The tasks used in the study had minimal mismatch compared to driving simulations. This result was explained by the observation that VR experiences are more novel and unique to older individuals, resulting in higher presence and less cybersickness. Increased presence is thought to reduce cybersickness by moving the user's attention away from the sensory mismatch they feel [78]. These mixed findings suggest that although age might play a role in cybersickness susceptibility, it may interact with related characteristics that come with age and can vary according to the intensity of the experience.

### 2.2 Field dependence/independence

The field dependence/independence construct was initially conceptualized by Witkin et al. [83] to describe an individual's tendency to rely more on visual or gravitational cues. It was found to be related to various other individual aspects, such as intelligence [66], academic achievement [54], and social behavior [67]. For this reason, it is sometimes described as a broader cognitive style related to the extent an individual's perception can be restructured to meet the task requirement [82]. However, there is still discussion to which extent field dependence/independence measures an ability or a broader thinking style [87].

The standard tests to measure field dependence are the Rod and Frame Test (RFT) and the Group Embedded Figures Test. The original RFT asks participants to vertically align a rod enclosed by a tilted frame. Field-independent individuals are able to align the rod to the true vertical, while FD individuals tend to be more strongly influenced by the frame tilt [81]. The Group Embedded Figures test asks participants to find a simple figure inside a complex one. The test requires completing as many tasks as possible in the allotted time. It was found that field-independent individuals also score better on this task [84]. The correlation between both tests led to the hypothesis that they measure the same trait: the degree to which an individual is influenced by the visual field.

#### 2.2.1 Cybersickness and Field Dependence/Independence

In one of his early studies, Witikins [80] noticed a connection between visually-induced motion sickness and field dependence. Studies using simulators found similar results. Using an automobile simulator, Barret and Thornton [9] found that field-independent participants had higher discomfort and illness, and supported fewer trials in the simulator. However, subsequent studies using a car simulator and a "haunted swing" (fixed swing and a movable simulated brick wall) found that field-dependent participants had the most discomfort [8, 10]. Studies looking at susceptibility questionnaires and field dependence measured by either the RFT or GEFT could not find a consistent link [17, 47, 85].

Similarly, studies in VR looking at the correlation between field dependence and cybersickness have yielded mixed results. Two consecutive studies by Manuvrier et al. found no significant correlation between virtual Rod and Frame test results and cybersickness [41, 42]. In both studies, participants used teleportation to travel in the virtual environment. More similar to ours, a more recent study [43] used a smooth and slow linear motion as stimuli. They found that the scores from the Rod and Frame test before immersion explained 26% of the variance in cybersickness. Unfortunately, they did not investigate the effect of sex, so it was not possible to compare the findings to a well-known predictor [39, 50]. In addition to considering sex, we used a drop-out paradigm to directly model the participant's study lifetime and used a more detailed model of the frame effect.

### 2.3 Motivation and Hypothesis

The goal of this study was to investigate how much cybersickness variability can be explained by field dependence. Identifying a significant predictor of cybersickness can improve the understanding of individual differences and enable better evaluations of mitigation techniques and predictive models [7].

Since individuals seem to adapt to become more field-independent following immersion [21, 77], we hypothesized that higher field dependence would be correlated with higher sickness levels and shorter times in the VR roller coaster simulation.

## 3 Methods

A challenge with cybersickness studies is how to deal with the inevitability of participant discontinuation [39]. Because of the wide range of individual susceptibility, any fixed stimulus will be too weak for some participants (limiting the ability to investigate sickness) while being excessive for others (exposing them to more discomfort than necessary). In addition, removing the participants who did not complete the study leaves behind valuable data, and analyzing them separately provides an incomplete view of the phenomenon.

To investigate our hypothesis, we conducted a time-to-event study. In this paradigm, we presented a cybersickness-inducing stimulus for 30 minutes or until a dropout event occurred. This allows us to naturally consider censoring, while achieving higher statistical power by considering if and when a participant drops out. It also allows us to obtain an interpretable effect sizes (hazard ratio).

Compared to sickness questionnaires, the time-to-event paradigm removes the scale ambiguity where the response may differ between participants experiencing the same symptoms, depending on how they map the scale. It also removes interpretation variability from questions with terms that can lead to different interpretations (e.g., fullness of head), leading to further variability [11]. Finally, because dropout is the final outcome, predicting it is more useful than intermediate measures, as the model will consider the combined effect of all variables that lead to dropout.

### 3.1 Ethics

The experimental protocol was approved by Northeastern University Institutional Review Board (#23-04-17). Informed consent was obtained prior to the start of the study. Participants in the study had the chance to receive a $50 gift card (odds 1:9).

### 3.2 VR Roller Coaster Stimulus

Similar to previous studies, we used a VR roller coaster as the stimulus [25, 69]. The one used in this study included a variety of unpredictable rotations and translations along different axes, as well as segments with low movement. The visual consists of a medieval desert city [2], combining European and Middle-Eastern styles as shown on Figure1. The buildings and the ground are mostly uniform in color with no textures. There is a single parallel source of light. For the path, a few decisions were made with the goal of increasing cybersickness: the path was designed around the buildings and close to the ground to increase optical flow; the track and the train were not visible to remove anticipation and fixed points of reference, respectively [60]; and participants did not have control of their movement (but could turn their head in any direction to look at the scene). The roller coaster project and source code are available at: https://github.com/realitydesignlab/cybersickness-coaster.

To define the acceleration levels, we conducted an informal evaluation with 21 participants. Each participant tried three versions of the roller coaster: gravity-like accelerations, unnatural accelerations, and a mix of the two. Gravity-like accelerations are consistent with physics, while unnatural ones go against gravity and inertia. The three variations followed the same path and were counterbalanced between participants. After each ride, participants completed a version of the VRSQ (without Q.6– fullness of head). After the last ride, participants were asked to rank the rides regarding comfort. Eight participants did not complete all versions and were excluded. There was no significant difference between the rides (non-parametric multiple comparisons, Fisher Method, 95% CI, p=.79). The version with unnatural accelerations was ranked as causing more cybersickness and was selected for this study.

The final roller coaster ride consisted of five sessions with three roller coaster rides per session, amounting to a maximum of 15 rides. A single ride lasted two minutes, and the same ride was used for the entire simulation to remain consistent. While on the roller coaster, we asked participants to search for green balloons placed along the track. Participants could pop them by looking in their direction (Figure 2). This minor task was to ensure participants would keep their eyes open and focused on the ride. It also encouraged participants to move their heads, which has been shown to increase cybersickness [24, 43]. Participants experienced the roller coaster sitting down in an office chair.

#### 3.2.1 Apparatus and Study Settings

The experimental application was developed using Unity 3D engine 2021.3.15f1 and presented in the HTC VIVE Pro Eye VR headset (1440 x 1600 resolution per eye, 98 degrees of horizontal field of view, 90 Hz refresh rate) connected to a desktop (Intel Core i5, NVIDIA GeForce RTX 4070 12Gb, 23Gb RAM). Participants interacted using the Vive standard controllers. For answering the cybersickness questionnaires, they used a pointing ray and the trigger

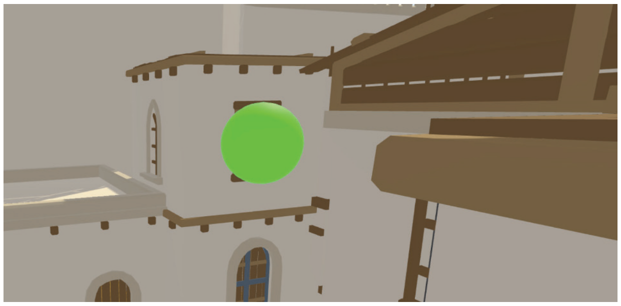

Figure 2: Green balloon that appeared in selected spots randomly throughout the ride.

to select. To adjust the Rod & Frame, they used the touchpad to rotate clockwise or counterclockwise and the trigger to select.

### 3.3 Measures

#### 3.3.1 Background Questionnaire

Participants completed an online background questionnaire to collect data on age, sex, experience with VR, general susceptibility to travel (car, bus, boat, and airplane), and roller coaster motion sickness.

#### 3.3.2 Random Dot Stereogram Test

Participants completed a random dot stereogram to assess for normal stereo vision. The static stereogram image was displayed on the headset containing the letters ”3D”. Participants were asked to describe what they saw. The experimenter gave time and adjusted the headset to ensure that the failure to describe the characters was not due to a bad headset fit.

#### 3.3.3 Time

For each participant, we measured the total time within the roller coaster. The time was tracked by the software application and manually terminated by the experimenter for participants who left in the middle of a ride. This was used as a measure of dropout.

#### 3.3.4 Cybersickness Questionnaire

Cybersickness is most commonly measured by the Simulator Sickness Questionnaire (SSQ), which has 16 questions in a 4-point Likert Scale [31]. A shorter, widely used questionnaire based on SSQ is the Virtual Reality Sickness Questionnaire (VRSQ) [34]. The VRSQ is highly correlated with SSQ but has better psychometric qualities to assess cybersickness because the components not related to VR were removed [68]. Since it is also shorter, it is more useful for quick assessment of cybersickness [20].

In the VRSQ development process, the SSQ nausea component was removed because it contributed less than the oculomotor and disorientation components and was mixed with other physiological reactions. However, we observed in our pilot study that nausea was a common reaction in the rollercoaster. On the other hand, participants were confused by the term ’fullness of head’. In our modified VRSQ, ’fullness of head’ was replaced with ’nausea’. Participants completed the questionnaire in VR using the controller before the first ride of the session and after each ride (Figure 3).

#### 3.3.5 Group Embedded Figures Test (GEFT)

The GEFT is a well-known measure of field dependence/independence [84]. The task consists of locating and tracing simple figures embedded in each of the 18 complex figures. The GEFT consists of two blocks of 5 minutes each. Scores range from 0 to 18. Individuals with scores lower than 12 are usually

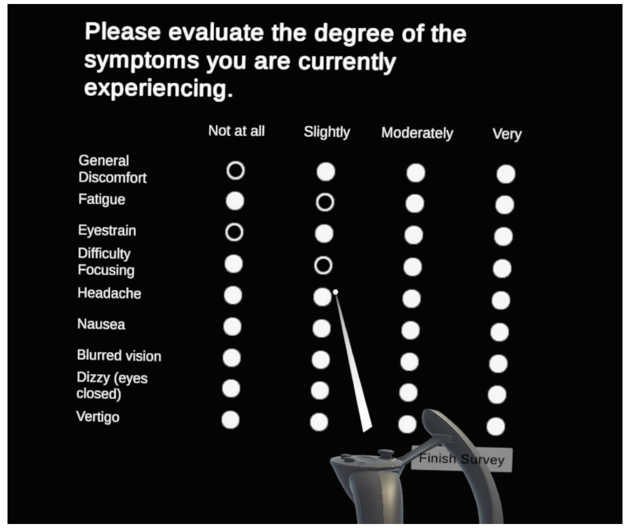


Figure 3: Screenshot of the sickness questionnaire completed before and after each ride.

Figure 4: Furnished Room version of the Rod and Frame test into VR. The two black spheres were always 180° from another and rotated as one.

considered field-dependent. In this study, we used the online version from Mind Garden [1].

#### 3.3.6 Rod and Frame Tests (RFT)

A version of the Rod and Frame test was developed to measure (static) field dependence in VR. The test consists in aligning a rod to the real vertical while in a rotated room [81]. Individuals who can adjust the rod to real vertical with less angle error are considered to rely more on internal bodily cues, as the visual rotation of the room has less influence on them, indicating higher field independence. Conversely, field-dependent individuals tend to make greater errors, as they rely more on the visual cues of the room when adjusting the rod. Instead of using a full rod, we followed the improved version of the Computerised Rod and Frame Test suggested by Docherty and Bagust [19], which reduces some prior observed bias from rendering on computer screens by showing only the edges of the rod using two disks.

Participants completed two instances of the RFT, one in which the room was empty (Figure 5) and one with objects (Figure 4). The empty layout was designed to more closely match the Portable Rod & Frame Test [53]. The furnished room was designed to look like a

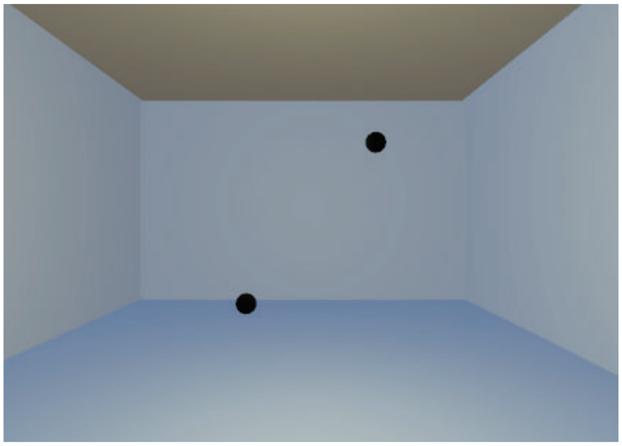

Figure 5: Empty Room.

household environment. The goal was to evaluate whether the RFT could be conducted in environments more likely to be found in a game or VR experience. We hypothesized that the additional cues in the furnished room would more strongly bias FD participants toward the frame direction instead of the real vertical.

The instructions given to the participant were to move the dots using the VIVE controller touchpad so that the spheres would appear vertically aligned with gravity. When the participants were satisfied with their alignment, they pressed the trigger on the controller to move to the next trial. Haptic feedback on the controller was used to keep the user's head and posture vertically still. The tests were conducted with the participant standing. Each participant completed 20 trials for each room with randomly rotated room frames and spheres (10 trials has been considered an adequate number in clinical studies of visual verticality [59]). The first five trials were excluded from the analysis to account for a learning effect, as some participants initially struggled with using the controller and misunderstood the task. In these first trials, some participants aligned the rod with the tilted room rather than the true vertical. This adjustment ensured that only data reflecting accurate task comprehension and controller usage were included in the final analysis.

### 3.4 Participants

Fifty participants were recruited through flyers on a large university campus. Participants were required to be between 18 and 35 years old, be able to follow instructions in English, have normal or corrected vision, and have no known vestibular or neurological condition. No participant reported a history of vestibular or neurological dysfunction. One individual was excluded for failing the stereovision test. After the study, three individuals were removed due to incomplete data (time data, RF data, GEFT data). Out of the valid 46 participants, one was removed because the measurements from the Rod & Frame test were larger than the $3^{rd}$ sample quartile plus 6 times the inter-quartile range.

The final sample consisted of $n = 45$ individuals, all students (18 females, 27 males, mean age = 22.9). Most participants did not play VR regularly, with 38.3% of the participants having never played VR, 42.6% having only tried VR a few times, and 19.1% playing a few hours every week. Most participants were not susceptible to getting sick from transportation and roller coasters, with 70% reporting rarely or never getting sick traveling by car, bus, boat, or airplane, and 68% reporting rarely or never getting sick on a roller coaster. Of those reporting sometimes, often, or always getting sick from transportation 64% were female, and for roller coasters, it was 58% female.

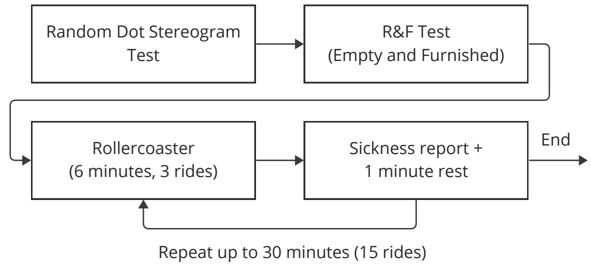


Figure 6: Outline of the study procedure

Table 1: Descriptive Statistics

| Measure | Min | Median | Mean | Max | SD |
|---|---|---|---|---|---|
| Age (years) | 18 | 23 | 22 | 32 | 3.45 |
| Sleep (1-10 quality) | 1 | 7 | 7.19 | 10 | 1.54 |
| GEFT (0-18 score) | 4 | 13 | 13.87 | 18 | 3.75 |
| Time (seconds) | 74 | 379 | 674.2 | 1894 | 613.9 |
| Raw RFT Empty (deg) | -5.59 | 0.10 | 0.18 | 6.91 | 2.57 |
| Raw RFT Furni. (deg) | -10.04 | -0.56 | -0.49 | 7.65 | 3.28 |
| Abs RFT Empty (deg) | 0.94 | 2.82 | 3.37 | 7.32 | 1.68 |
| Abs RFT Furni. (deg) | 1.19 | 3.77 | 4.35 | 16.13 | 2.85 |
| Sickness Score (0-27) | 0 | 6 | 7.96 | 19 | 5.50 |

### 3.5 Procedure

Before the in-person part of the study, a copy of the consent form was sent to participants, along with links to the online assessment of the Group Embedded Figures Test (GEFT) and the Background Questionnaire. When participants arrived, we obtained written consent and asked additional questions about their sleep hours and quality. After headset fitting, interpupillary distance, and eye tracking calibration, participants completed the stereogram test. Following, they were introduced to the Rod and Frame Test controls, and had some time to get used to them. Participants were instructed to use the controller's touchpad to align the dots so that they would appear vertically aligned with respect to gravity. Next, they completed the two versions of the RFT, one empty room and one furnished room. The order of the rooms was counterbalanced.

After the initial preparations, participants started the roller coaster rides. After each session (three rides), participants were given a one-minute break to rest without the headset. The participants continued this procedure until they felt too sick to continue or made it through all 15 rides (Figure 6). Before the first ride and after each ride, participants completed the modified VRSQ questionnaire. The questionnaire took around 15 seconds to complete. Participants were reminded of the option to withdraw at any point and encouraged to end the study if they felt too nauseous or uncomfortable to continue any further. The in-person study took roughly one hour.

## 4 Results

The data collected was analyzed using R version 4.3.2 [71]. Survival analysis was conducted using the robust version of the Cox Proportional Hazards model from the Survival package [72] version 3.5.7. Non-parametric multiple comparison tests used the nparcomp package version 3.0 [36]. Plots used ggplot2 [79] version 3.5.1. For all tests, the statistical significance was set at 5%. The descriptive statistics for the main variables in the study are listed in Table 1.

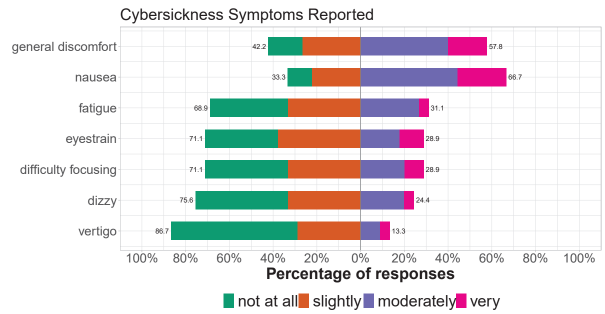


Figure 7: Proportion of Cybersickness Symptoms with Highest Values

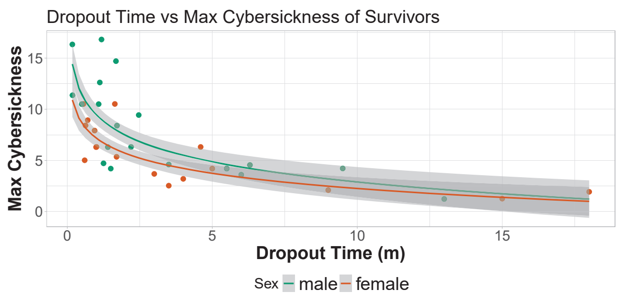


Figure 8: Maximum cybersickness score by dropout time. Time to dropout reduced exponentially with the maximum experienced sickness level.

### 4.1 Cybersickness Impact

We first report the results from the cybersickness questionnaire applied at every 3 rides. Among the symptoms listed, nausea was the most common symptom, felt moderately or very strongly by most participants (66.7%, *n*=30). Symptoms of fatigue (31.1%, *n*=14), eyestrain (28.9%, *n*=13), difficulty focusing (28.9%, *n*=13), dizziness (24.4%, *n*=11), and vertigo (13.3%, *n*=6) were slightly or not felt by most participants. Most participants indicated moderate or very strong general discomfort (57.8%, *n*=26).

At the highest values reported (Figure 7), a multiple comparison test procedure with Tukey's correction indicated that nausea was statistically higher than all other symptoms, excluding general discomfort (which would reasonably include nausea): than fatigue ($\Delta^p$ = -.187, p = 0.009), eyestrain ($\Delta^p$ = -.189, p=.009), difficulty focusing ($\Delta^p$ = -.206, p=.005), dizziness ($\Delta^p$ = -.245, p< .001), and vertigo ($\Delta^p$ = -.322, p< .001). Where $\Delta^p$ is the estimator for pairwise differences in pseudo-ranks.

The majority of participants were unable to finish all the rides. Only 8 individuals (18%) reached the end of the 30-minute experience, with 37 participants (82%) dropping out within 17 minutes of exposure. There were no technical errors or outside factors causing a participant to drop out. No emesis was observed.

As expected, participant dropout time was consistent with high levels of self-reported cybersickness, with all dropout participants expressing discomfort as the reason for ending the study. Eighteen participants had non-zero baseline cybersickness scores (min=1, median=2, max=8). Time to dropout increased exponentially with the decrease in cybersickness (Figure 8). A generalized linear model with a log link for cybersickness and sex indicated significance for both predictors, with the relative contribution of sex in the explained variance being very small (Table 2).

From the participants, field-independent males dropped at the slowest rate while field-dependent females dropped at the fastest

Table 2: Cybersickness and Dropout Time

| Variable | Estimate | Std. Error | p | Contribution |
|---|---|---|---|---|
| Intercept | 6.65 | .092 | $< .001$ | - |
| Cybersickness | -.20 | .043 | $< .001$ | .521 |
| SexFemale | -.034 | .127 | .010 | .009 |

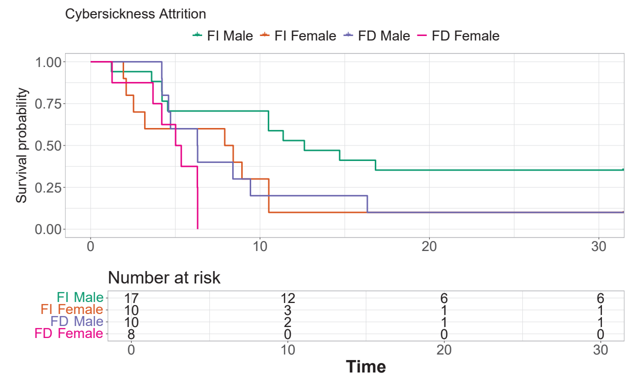


Figure 9: Kaplan-Meier curve of survival probability. For visualization, participants were split into field-dependent and independent by the mean.

rate. At the end of the study, only one female (5%) and seven males (39%) remained (Figure 9). A Wilcoxon rank sum test indicated no significant difference between the cybersickness scores of males and females participants ($W = 234.5, p = .853$).

When looking at the background questionnaire, there was no significant difference in the susceptibility questions (W = 130, p-value = 0.60) between the participants who stayed (min=5, mean = 6.75, max = 10) and the ones who dropped (min=2, mean = 7.97, max = 15).

### 4.2 Field Dependence Measurements

We evaluated three measurements for field dependence: GEFT, RFT in an empty room, and RFT in a furnished room. Figure 10 shows the distribution, and Table 1 shows the descriptive statistics.

Both versions of the RFT showed a strong frame effect. The mean absolute error for the empty and furnished rooms were 3.37 and 4.35 degrees respectively. The furnished room showed more extreme errors errors and a slightly higher mean. The difference was not significant for the raw errors ($t(44) = 1.24, p = .22$) but was significant for the absolute errors ($t(44) = -2.8, p = .007$). The mean error at zero and 90 frame tilt was similar for the empty (1.23 degrees) and the furnished rooms (1.23 degrees).

The Spearman's correlation between GEFT and the Rod and Frame tests was weak. However, it was similar to the ones reported by Panek et al. [56] for the mean age of 21.2 years ($\rho = -0.52$). The correlation was significant for the absolute angle error for the empty room and GEFT ($\rho = 0.56, p < 0.001$). The correlation for the furnished room and GEFT was lower and not significant ($\rho = -0.28, p = 0.058$). The angular errors for the two Rod & Frame versions were moderately correlated ($\rho = 0.364, p = 0.014$).

A Welch t-test did not show a statistical difference between males and females for the raw scores of the RF test in the empty room (t(42)=-0.93, p=.35), furnished room (t(34)=-0.013, p=0.99), neither for the absolute errors in the empty room (t(37)=0.13, p=0.89) or furnished room (t(41)=-0.24, p=0.81). A Wilcox signed rank test also did not indicate a significant difference between males and females for the GEFT (W=190, p=.221). The effect size was negligible (r = -0.03).

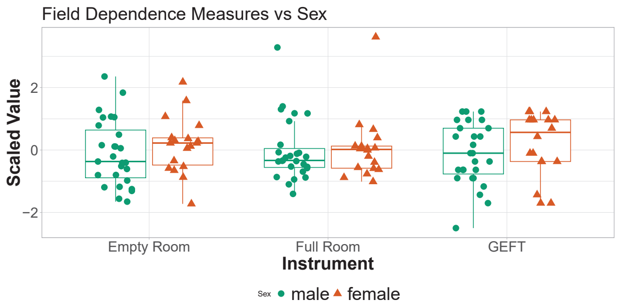


Figure 10: Field Dependence Measurements

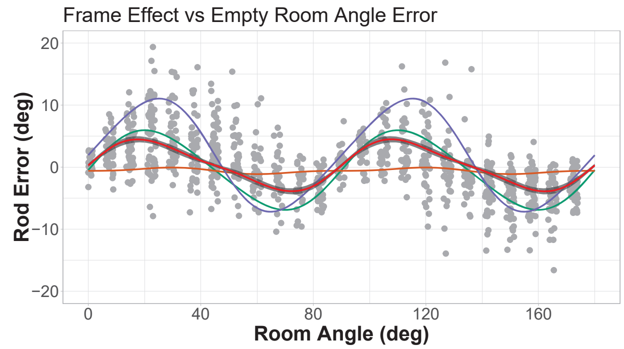


Figure 11: Rod & Frame Fit. The lines without bands display the sample fit for 3 participants. The overall fit for all participants is shown in red, along with a 95% confidence interval band.

To compare with previous studies, we tested the correlation between sickness scores and the average rod errors across all trials. The correlation was not significant (p=.892, $\rho$=-.021). We also conducted a simple linear regression analysis ($VRS \sim Abs(RF_{Empty})$) and did not find statistical significance ($R^2$ = -0.023, F(1,43)=0.004, p=0.949).

### 4.3 Rod and Frame Scoring

In order to obtain a single estimate of the frame effect per participant, the trials need to be aggregated in some way. A common scoring method is to use the mean of absolute errors [6] or a weighted mean based on the direction of the frame influence (towards or opposite to the frame tilt). Comparisons of absolute and algebraic scoring systems [4, 65] indicate that the mean absolute error has reasonable reliability and correlation with multiple measures of spatial ability. This metric combines all errors regardless of the direction and cause.

For this research, we first modeled the rod and frame responses based on Mittelstaedt's model for the Subjective Visual Vertical (SVV) [48]. Mittelstaedt's original model includes three components: an estimate of the gravity vector from the otolith organs, a bias vector aligned with the long axis of the body (idiotropic vector), and a visual frame vector that includes the cues from the environment. The frame vector introduces periodic deviations that can be modeled as a harmonic sum, whose terms match the periodicities of the visual components of the environment. We used the harmonic fit used by Vingerhoets et al. [74] which consists of two terms with periods of 90° and 180° :

$$SVV = F_0 + F_1 \cdot sin(2\theta_w - \Delta\phi_1) + F_2 \cdot sin(4\theta_w - \Delta\phi_2)$$

In the model, $\theta_w$ is the frame tilt angle in world coordinates, $F_0$ is a constant bias across all angles, $F_1$ and $F_2$ are the coefficients for each term, and $\Delta\phi_1$ and $\Delta\phi_2$ represent the corresponding phase shift for each harmonic component. This model thus captures 90° 4-fold error symmetries of a square frame and 180° up/down error symmetries related to gravity and idiotropic vector.

Following Vingerhoets procedure, each participant's Rod and Frame deviation from vertical was fit, and the peak amplitude of the fit model (SVV) for each person was used as a score (these correspond to the points where the sum of the three terms is maximal). Head tilt was constrained by our design and we assumed no phase lag between the frame rotation and the corresponding SVV error cycle, making $\Delta\phi = 0$. The resulting fit can be seen on Figure 11. The model with both harmonic terms was significantly better than a reduced model with just the first term and without the *sin* term ($F = 15.45, p < 0.001$). The complete model also had the lower AIC. After normalization, the amplitude statistics for the empty room ($RF_{\text{Empty}}$) were min = -1.72, median = 0.06, max=2.36; and for the furnished room ($RF_{\text{Empty}}$) min=-1.40, median=-0.22, max=3.62).

The Spearman's rank correlation between the two versions of the rod and frame was stronger for the amplitude ($\rho = 0.646, p < 0.001$) than the raw angular errors ($\rho = 0.364, p = 0.0145$). The correlation between the peak amplitude and GEFT for the empty room was stronger and significant ($\rho$=-.350, p=.018). The correlation between GEFT and amplitude of the empty room was also stronger but did not reach significance ($\rho$=-.284, p=0.058). Finally, dropout time was significantly correlated with the Amplitude (p=.026, $\rho$=-.331).

### 4.4 Survival Analysis

Survival analysis refers to a family of statistical techniques designed to analyze time-to-event. In this research, the event is the participant dropping the simulation and time-to-event the duration of the experience.

The Kaplan-Meier plot is a non-parametric statistic used to estimate the survival function from time-to-event data. It can provide a useful visualization of the raw cumulative survival curve, plotting the probability of survival against time. Because it is restricted to categorical variables, we discretized the GEFT scores by the mean for an initial overview of the results (Figure 9). For the subsequent analysis we use the values without discretizing them into groups.

Unfortunately, besides requiring categorical independent variables, the Kaplan-Meier statistic does not provide an estimate of relative risk, and does not allow the adjustment of confounders [3]. The Cox regression model is a semiparametric model to estimate the hazard rate, i.e., the ratio between the hazards of two groups of individuals and its change along a continuous variable.

For time-to-event outcomes prediction models, an established rule of thumb for the required sample size is to ensure at least 10 events for each predictor parameter. Although this rule has been found to be conservative [75], we limited our model to only three parameters, resulting in 12.3 events per parameter.

#### 4.4.1 Model Selection

We evaluated models containing our variable of interest (FD) and other known predictors of cybersickness: age, sex, and sleep. For field dependence, we compared models with GEFT and both versions of the Rod and Frame test (empty and furnished room). all variables were normalized. Table 3 lists the AIC values and Nagelkerke's pseudo-$R^2$ for the tested models. The pseudo-$R^2$ reflects how well the model explains the variability in survival times (the higher the better), and the AIC how well it fits the data while penalizing for model complexity (the lower the better). The model Surv $\sim$ $\text{RF}_{\text{Empty}}$ + Sex was selected since it had the lowest AIC and the highest $R^2$. Variance inflation factors indicated no significant collinearity between Sex and $RF_{Empty}$ (1.036).

Table 3: Akaike Information Criteria and Nagelkerke pseudo-$R^2$. Sex and RF scores in the empty room showed best explanatory power (model 2). *RF* refers to the SVV amplitudes of each room.

| | Model | AIC | Nagelkerke |
|---|---|---|---|
| 1 | Surv $\sim$ GEFT + Sex | 235.1 | 0.124 |
| 2 | Surv $\sim$ $RF_{Empty}$ + Sex | $\boxed{231.6}$ | $\boxed{0.189}$ |
| 3 | Surv $\sim$ $RF_{Furni}$ + Sex | 235.1 | 0.125 |
| 4 | Surv $\sim$ GEFT + Sleep | 240.5 | 0.012 |
| 5 | Surv $\sim$ $RF_{Empty}$ + Sleep | 236.8 | 0.089 |
| 6 | Surv $\sim$ $RF_{Furni}$ + Sleep | 240.5 | 0.012 |
| 7 | Surv $\sim$ GEFT + Age | 240.1 | 0.001 |
| 8 | Surv $\sim$ $RF_{Empty}$ + Age | 236.6 | 0.094 |
| 9 | Surv $\sim$ $RF_{Furni}$ + Age | 241.0 | 0.002 |

Table 4: Cox Proportional Hazards model for Sex and RF in the empty room

| Variable | Hazard Ratio | 95% Confidence Interval | p |
|---|---|---|---|
| Amplitude | 1.37 | 1.047 - 1.790 | 0.022 |
| SexFemale | 2.26 | 1.105 - 4.606 | 0.025 |

The partial likelihood ratio test against the null model (intercept only) was significant ($\chi^2$ = 9.39, p = .009). Sex was a significant predictor, with females associated with a risk increase of 2.26 or 126%. The score from the Rod and Frame in the empty room was also significant. Individuals lying more on the field-dependent side of the spectrum - those who made larger errors when aligning the rod to the real vertical - showed an increased risk of dropout of 37%.

Median time-to-event was 9.4 minutes. As a measure of effect, the model's predicted median time-to-event for females was 6.3 min. and 11.4 min. for males. One degree of error amplitude reduces female median time by 0.42 minute and male time by 1 minute. $RF_{Empty}$ raw amplitudes ranged from 0.53 to 9.12, with median 4.28 and mean 4.16.

## 5 Discussion

The aim of this study was to investigate whether field dependence is useful as a predictor of cybersickness. We discuss here the main findings and their relation to earlier studies.

### 5.1 Cybersickness Impact

The roller-coaster design was highly effective in quickly inducing cybersickness, leading to participants' dropout, mostly due to moderate to intense nausea. The analysis of drop-out time and self-reported cybersickness showed that the increase in cybersickness score led to exponentially less time on the rollercoaster. This correlation has also been found in prior work [7, 57].

Participants in the study dropped out at different levels, which indicates the high subjectivity of the questionnaire scores. Given a specific cybersickness score, it is difficult to predict how much time an individual can still tolerate before having to leave. Unfortunately, the decision to terminate the study can also be affected by factors not directly related to the feeling of being sick, such as motivation and how participants see the experience.

When looking at the cybersickness scores, it is important to note that our study applied the cybersickness questionnaire multiple times to allow us to verify the effect of our manipulation over time and establish an association with dropout time. However, Young et al. showed that applying the SSQ multiple times can increase the scores [86]. Although others were not able to replicate such effects in SSQ or FMS [14, 32], our score values may not be directly comparable to studies with a single application.

Surprisingly, a few participants stayed until the end of the study and reported low levels of cybersickness during the entire time. Some of them reported that they became used to the experience, and one expressed that they could stay there "for the whole day". In some cases, the scores decreased over time. These findings agree with the ones reported by Park et al. on car simulator dropouts [57]. The fact that we found no significant difference in the susceptibility questions indicates that our model was more sensitive.

Besides reduced vestibular sensitivity, behavioral, physiological, and cognitive adaptation may explain some of the observed resilience [76]. For example, learning the dynamics of the virtual experience and adapting their internal model may reduce the sensory conflicts [13]. Field-dependent participants have also been observed to switch to a more field-independent mode after VR experience, particularly in those with lower presence and a high level of cybersickness [21, 42].

### 5.2 Field Dependence Measures

The study's results show a strong frame effect and align well with the predictions from Mittelstaedt's model [49]. Outside the neutral positions, the rotated room induced a periodic modulation with positive peaks close to 27 and 117 degrees [74].

Our study evaluated two different environments as a frame: an empty blank room and a furnished room. One of our hypotheses was that the furnished room would be similar enough to the empty room. In fact, we expected that the extra cues would more strongly bias field-dependent participants towards the frame direction. As expected, the furnished room induced more extreme errors, higher variance, and a slightly higher mean error. However, the furnished room performed slightly worse as a predictor, which was unexpected. One possible reason is that the harmonic fit we used only had two terms and may not have been sufficient to capture the effect of the different visual cues in the room, resulting in a poor fit.

Regarding the GEFT test, we found that the correlation was small but agreed with the previous study on EFT (Embedded Figures Test), developed by Witikin prior to GEFT. The correlation between RFT and ETF is 36% to 74%. In our case the correlation with GEFT was 56% (Section 4.2). As mentioned by Barret [8], the RFT is extremely reliable, while the ETF not as much. The fact that the GEFT score had lower prediction power than the Rod and Frame in the empty room is evidence to support that these instruments measure different dimensions, and why GEFT was not very useful to explain the dropout time.

### 5.3 Field Dependence as a Predictor to Cybersickness

While many studies report the number of participant dropouts [15, 39], there are very few studies using a time-to-event paradigm. We chose this method to obtain a true marker of the participant's maximum sickness level and to address the inevitability of participant discontinuation.

Using data from a non-VR driving simulation study, Matas et al. [44] developed a CoxPH model to predict the dropout stage in a driving simulation (in blocks of 10 minutes each). Their model included as independent variables age, gender, motion sickness, medical conditions, and MMSE scores (Mini-Mental State Examination), a short questionnaire designed to identify possible dementia or mild cognitive impairments. Although they did not discuss the fitness of this model, their hazard ratio for gender (1.94) was similar to the 2.26 ratio found in this study (Table 4).

As mentioned before, only one study was able to demonstrate some link between field dependence and cybersickness. Manuvrier et al. [43] found that field dependence explained 26.9% of the cybersickness variance. We failed to replicate their results by running a simple linear regression (Section 4.2). [43] This result was also not replicated by a more recent work [28]. Josupeit evaluated four metrics: the mean absolute deviation from the compatible and incompatible trials, the average of all trials but the neutral ones, subtracting the constant error from the mean right-tilted frame condition, and subtracting the mean angle error in the neutral condition from the mean angle error of the compatible and incompatible conditions. None of the metrics were correlated with the scores of the Rod and Frame administered prior to the experience. Similarly, we did not find a significant correlation between sickness and the average of all trials (Section 4.2).

Compared to prior work, our approach differs in operationalizing field dependence and how we use it to predict the participant experience. The challenge with previous approaches is that the frame's influence will vary with the individual and frame tilt, but will be centered around zero if the participants keep their head aligned with the vertical. Another issue is that although participants dropped the study due to cybersickness discomfort, individuals anchor the scale differently. For example, a study with a termination cut-out score will still subject the participants to different levels of discomfort. In our study, although dropout time was not correlated with sickness scores, it was significantly correlated with the Amplitude measure we adopted (Section 4.3)..

Many studies have tried to investigate field dependence as a sole predictor. This is very unlikely to be true, and misspecifying the model by not including known effects can lead to erroneous conclusions. However, we think that a prediction model is still useful even if there are mediators and other variables unaccounted. In fact, a small number of easy-to-obtain predictors with a reasonable explanatory power would be the extremely useful.

In our study, each 1-unit increase in the field dependence score, increases the risk of dropout by 37%, so the impact of sex is equivalent to approximately 2.59 units increase in field dependence or 1.23 standard deviation. The contribution to the increase in the hazard rate from field dependence was $\sim$ 1.6 times smaller than sex (Table 2). The general direction of the sex effect was concordant with the literature: participants assigned as female at birth had a higher risk of dropping out. The other two variables investigated, age and sleep, were not better predictors of dropout risk than sex. Age was likely uninformative given that our sample had a limited range (18 to 32 years old) and the model had the same explanatory power as one with sleep. The fact that both field dependence and sex were significant predictors of dropout suggests that these two factors have somewhat different contributions.

It is interesting to note that although sex was a significant predictor in the survival analysis, a straightforward Wilcoxon Rank Sum test for sex differences on cybersickness scores was inconclusive. Assuming there is a true difference, a possible explanation is that our sex unbalanced design reduced the power of the test. Another possibility is that the true difference was confounded by the effect of field dependence. If true, this last hypothesis may explain some of the contradictory findings regarding sex differences and/or provide a more mechanistic explanation for them.

## 6 Limitations and Future Work

In our study, participants were in an upright position. According to prior work, the frame effect is more pronounced at a 60-degree head tilt. This protocol could strengthen the power to differentiate individuals with fewer trials. In this work, we proposed using the amplitude of the fit of Vingerhoets' et al. harmonic model. However, the authors suggest that a Bayesian model can provide a better fit for smaller tilt angles [74].

A study published while the current study was underway indicates that the experimenter's gender can affect the result of cybersickness reporting. Maneuvrier observed that women evaluated by men seemed to report more cybersickness than men evaluated by men, while women and men evaluated by women reported similar levels [40]. We had a single male experimenter, and although we did not find a statistical difference, varying the experimenter's gender seems appropriate for future work, so the finding can be replicated and also investigated for the rod and frame scores. Gender and sex differences regarding cybersickness continue to be a complex issue, and more studies are needed to understand the role of field-dependence and sex in cybersickness. We also reinforce the call for the identification of other variables and for studies with higher explanatory power [30]. Our model only included two variables, since we wanted to avoid bias and have reliable confidence interval coverage. Larger sample sizes would allow models with more variables and interactions, which may help to improve model fit and thus better prediction of dropout times.

The roller coaster used in this study was designed to elicit strong discomfort quickly. With the exception of a few games, our stimulus was unlike most everyday VR experiences, which are carefully designed for comfort. More comfortable experiences lead to a slow buildup of sickness, until users feel the need to stop. Prior work has observed that VR immersion can improve the performance of field-dependent individuals and lead to different types of adaptation [21, 76]. Further studies are needed to understand if the results we found still apply over slower increases in sickness.

Cybersickness studies using time to drop out as the main metric will have limited ability to identify differences if they fail to identify participants who dropped out for reasons other than cybersickness. In our study, we interviewed the participants afterwards to understand the reason. We also observed that the maximum reported cybersickness reduced exponentially with participants' time in the experiment. However, we cannot rule out that some participants may still have opted to terminate the study for other reasons while reporting the main cause to be cybersickness.

Finally, our study used a slightly modified version of VRSQ. Although it was sufficient to monitor the sickness effects over time and the prevalence of symptoms, we advise against directly comparing our scores with the validated standard version. In addition, even though it was clear from participants' comments that sickness was the major reason for discontinuation, future studies using this paradigm would benefit from systematically capturing the reason for dropout.

## 7 Conclusion

Our findings show that field dependence, as measured by the Rod and Frame test, can be helpful as a predictor of cybersickness dropout. The underlying frame effect was observed both in empty and furnished test environments, although the former had more explanatory power in the hazard model. On the other hand, the cybersickness questionnaire was not helpful in predicting dropout.

We found that while field dependence was statistically significant as a predictor, the risk associated with females was much greater (126%) than the risk due to being field-dependent (37%). That said, in our sample, field dependence was still better than age or sleep quality as a predictor. Regarding the measurement of field dependence, we could not replicate previous findings on the usefulness of the absolute means of the frame errors. However, we proposed a new metric based on the amplitude of Mittelstaedt's harmonic fit that was effective for this end. These findings may help to elucidate the individual differences behind cybersickness and to advance the design of methods for assessing them.

## 8 Acknowledgements

We would like to thank Joe Conwell for the initial work on the implementation and Todd Oggle for the initial discussion on field dependence.

## References

[1] Mind garden inc. https://www.mindgarden.com/. Acessed: 2025-04-01.

[2] Polylised - medieval desert city. https://assetstore.unity.com/packages/3d/environments/historic/polylised-medieval-desert-city-94557. Acessed: 2025-04-01.

[3] S. Abd ElHafeez, G. D'Arrigo, D. Leonardis, M. Fusaro, G. Tripepi, and S. Roumeliotis. Methods to analyze time-to-event data: The cox regression analysis. *Oxidative medicine and cellular longevity*, 2021(1):1302811, 2021.

[4] M. J. Allen, M. Garcia, and L. B. Bealessio. Measurement of rod-and-frame test performance. *Perceptual and Motor Skills*, 54(3):915–922, 1982.

[5] L. L. Arns and M. M. Cerney. The relationship between age and incidence of cybersickness among immersive environment users. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 267–268. IEEE, 2005.

[6] J. Bagust, S. Docherty, and R. A. Razzak. Rod and frame alignment times increase when the frame is tilted. *Psychology and Behavioral Sciences*, 2(2):66–72, 2013.

[7] S. A. Balk, M. A. Bertola, and V. W. Inman. Simulator sickness questionnaire: Twenty years later.(2013), 2013.

[8] G. V. Barrett. Relation between embedded figures test performance and simulator behavior. *The Journal of applied psychology*, 53(3):253–4, Jun 1969.

[9] G. V. Barrett and C. L. Thornton. Relationship between perceptual style and simulator sickness. *The Journal of applied psychology*, 52(4):304–8, Aug 1968.

[10] G. V. Barrett, C. L. Thornton, and P. A. Cabe. Cue conflict related to perceptual style. *The Journal of applied psychology*, 54(3):258–64, Jun 1970.

[11] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pp. 464–467. IEEE, 2020.

[12] J. Bos, R. Van Leeuwen, and T. Bruintjes. Motion sickness in motion: from carsickness to cybersickness. *Nederlands tijdschrift voor geneeskunde*, 162:D1760–D1760, 2018.

[13] J. E. Bos, W. Bles, and E. L. Groen. A theory on visually induced motion sickness. *Displays*, 29(2):47–57, 2008.

[14] P. Brown and W. Powell. Pre-exposure cybersickness assessment within a chronic pain population in virtual reality. *Frontiers in Virtual Reality*, 2:672245, 2021.

[15] E. Chang, H. T. Kim, and B. Yoo. Virtual reality sickness: a review of causes and measurements. *International Journal of Human–Computer Interaction*, 36(17):1658–1682, 2020.

[16] J. Clifton and S. Palmisano. Effects of steering locomotion and teleporting on cybersickness and presence in hmd-based virtual reality. *Virtual Reality*, 24(3):453–468, 2020.

[17] R. F. Deich and P. M. Hodges. Motion sickness, field dependence, and levels of development. *Perceptual and motor skills*, 36(3):1115–20, Jun 1973.

[18] A. T. Dilanchian, R. Andringa, and W. R. Boot. A pilot study exploring age differences in presence, workload, and cybersickness in the experience of immersive virtual reality environments. *Frontiers in Virtual Reality*, 2:736793, 2021.

[19] S. Docherty and J. Bagust. From line to dots: an improved computerised rod and frame system for testing subjective visual vertical and horizontal. *BMC research notes*, 3:9, 2010. doi: 10.1186/1756-0500-3-9

[20] J. Dong, K. Ota, and M. Dong. Why vr games sickness? an empirical study of capturing and analyzing vr games head movement dataset. *IEEE MultiMedia*, 29(2):74–82, 2022.

[21] L. Fantin, G. Ceyte, E. Maïni, G. Hossu, and H. Ceyte. Do individual constraints induce flexibility of visual field dependence following a virtual immersion? effects of perceptive style and cybersickness. *Virtual Reality*, 27(2):917–928, 2023.

[22] J. M. Fulvio, M. Ji, and B. Rokers. Variations in visual sensitivity predict motion sickness in virtual reality. *Entertainment Computing*, 38:100423, 2021.

[23] J. F. Golding. Motion sickness susceptibility. *Autonomic Neuroscience*, 129(1-2):67–76, 2006.

[24] P. Howarth and M. Finch. The nauseogenicity of two methods of navigating within a virtual environment. *Applied Ergonomics*, 30(1):39–45, 1999.

[25] R. Islam, Y. Lee, M. Jaloli, I. Muhammad, D. Zhu, and J. Quarles. Automatic detection of cybersickness from physiological signal in a virtual roller coaster simulation. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pp. 648–649. IEEE, 2020.

[26] A. Jasper, N. Cone, C. Meusel, M. Curtis, M. C. Dorneich, and S. B. Gilbert. Visually induced motion sickness susceptibility and recovery based on four mitigation techniques. *Frontiers in Virtual Reality*, 1:582108, 2020.

[27] A. Jasper, N. C. Sepich, S. B. Gilbert, J. W. Kelly, and M. C. Dorneich. Predicting cybersickness using individual and task characteristics. *Computers in Human Behavior*, 146:107800, 2023.

[28] J. Josupeit. In rod we trust–the evaluation of a virtual rod and frame test as a cybersickness screening instrument. *PloS one*, 19(11):e0313313, 2024.

[29] N. Kawano, K. Iwamoto, K. Ebe, B. Aleksic, A. Noda, H. Umegaki, M. Kuzuya, T. Iidaka, and N. Ozaki. Slower adaptation to driving simulator and simulator sickness in older adults aging clinical and experimental research. *Aging clinical and experimental research*, 24:285–289, 2012.

[30] J. W. Kelly, S. B. Gilbert, M. C. Dorneich, and K. A. Costabile. Gender differences in cybersickness: Clarifying confusion and identifying paths forward. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 283–288. IEEE, 2023.

[31] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[32] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Human factors*, 53(4):415–426, 2011.

[33] B. Keshavarz, R. Ramkhalawansingh, B. Haycock, S. Shahab, and J. Campos. Comparing simulator sickness in younger and older adults during simulated driving under different multisensory conditions. *Transportation research part F: traffic psychology and behaviour*, 54:47–62, 2018.

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

[35] M. Kim. Cybersickness: Why people experience motion sickness during virtual reality. *Inside Science*, August 14 2019.

[36] F. Konietschke, M. Placzek, F. Schaarschmidt, and L. A. Hothorn. nparcomp: An R software package for nonparametric multiple comparisons and simultaneous confidence intervals. *Journal of Statistical Software*, 64(9):1–17, 2015.

[37] J. J. LaViola Jr. A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin*, 32(1):47–56, 2000.

[38] T. Luong, A. Pléchata, M. Möbus, M. Atchapero, R. Böhm, G. Makransky, and C. Holz. Demographic and behavioral correlates of cybersickness: A large lab-in-the-field study of 837 participants. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 307–316. IEEE, 2022.

[39] C. MacArthur, A. Grinberg, D. Harley, and M. Hancock. You're making me sick: A systematic review of how virtual reality research considers gender & cybersickness. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pp. 1–15, 2021.

[40] A. Maneuvrier. Experimenter bias: exploring the interaction between participant's and investigator's gender/sex in vr. *Virtual Reality*, 28(2):96, 2024.

[41] A. Maneuvrier, L. M. Decker, H. Ceyte, P. Fleury, and P. Renaud. Presence promotes performance on a virtual spatial cognition task: Impact of human factors on virtual reality assessment. *Frontiers in Virtual Reality*, 1:571713, 2020.

[42] A. Maneuvrier, L. M. Decker, P. Renaud, G. Ceyte, and H. Ceyte. Field (in) dependence flexibility following a virtual immersion is associated with cybersickness and sense of presence. *Frontiers in Virtual Reality*, 2:706712, 2021.

[43] A. Maneuvrier, N.-D.-T. Nguyen, and P. Renaud. Predicting vr cybersickness and its impact on visuomotor performance using head rotations and field (in) dependence. *Frontiers in Virtual Reality*, 4:1307925, 2023.

[44] N. A. Matas, T. Nettelbeck, and N. R. Burns. Dropout during a driving simulator study: A survival analysis. *Journal of safety research*, 55:159–169, 2015.

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

[46] M. Melo, G. Gonçalves, D. Narciso, and M. Bessa. Impact of different role types and gender on presence and cybersickness in immersive virtual reality setups. In *2021 international conference on graphics and interaction (ICGI)*, pp. 1–8. IEEE, 2021.

[47] C. S. Mirabile, jr, B. C. Glueck, and C. F. Stroebel. Susceptibility to motion sickness and field dependence-independence as measured with the rod and frame test. *Neuropsychobiology*, 2(1):45–51, 1976.

[48] H. Mittelstaedt. The subjective vertical as a function of visual and extraretinal cues. *Acta psychologica*, 63(1):63–85, 1986.

[49] H. Mittelstaedt. The subjective vertical as a function of visual and extraretinal cues. *Acta psychologica*, 63(1-3):63–85, Dec 1986.

[50] J. M. Mittelstaedt. Individual predictors of the susceptibility for motion-related sickness: a systematic review. *Journal of Vestibular Research*, 30(3):165–193, 2020.

[51] J. Munafo, M. Diedrick, and T. A. Stoffregen. The virtual reality head-mounted display oculus rift induces motion sickness and is sexist in its effects. *Experimental brain research*, 235:889–901, 2017.

[52] Y. Nam, U. Hong, H. Chung, and S. R. Noh. Eye movement patterns reflecting cybersickness: evidence from different experience modes of a virtual reality game. *Cyberpsychology, Behavior, and Social Networking*, 25(2):135–139, 2022.

[53] P. K. Oltman. A portable rod-and-frame apparatus. *Perceptual and motor skills*, 26(2):503–6, Apr 1968.

[54] B. U. Onyekuru. Field dependence-field independence cognitive style, gender, career choice and academic achievement of secondary school students in emohua local government area of rivers state. *Journal of Education and Practice*, 6(10):76–85, 2015.

[55] S. Palmisano, R. S. Allison, and J. Kim. Cybersickness in head-mounted displays is caused by differences in the user's virtual and physical head pose. *Frontiers in Virtual Reality*, 1:587698, 2020.

[56] P. E. Panek, L. G. Funk, and P. K. Nelson. Reliability and validity of the group embedded figures test across the life span. *Perceptual and motor skills*, 50(3 Pt 2):1171–4, Jun 1980.

[57] G. D. Park, R. W. Allen, D. Fiorentino, T. J. Rosenthal, and M. L. Cook. Simulator sickness scores according to symptom susceptibility, age, and gender for an older driver assessment study. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 2702–2706. SAGE Publications Sage CA: Los Angeles, CA, 2006.

[58] K. Petri, K. Feuerstein, S. Folster, F. Bariszlovich, and K. Witte. Effects of age, gender, familiarity with the content, and exposure time on cybersickness in immersive head-mounted display based virtual reality. *American Journal of Biomedical Sciences*, 12(2), 2020.

[59] C. Piscicelli and D. Perennou. Visual verticality perception after stroke: A systematic review of methodological approaches and suggestions for standardization. *Annals of physical and rehabilitation medicine*, 60(3):208–216, 2017.

[60] T. Porcino, D. Trevisan, and E. Clua. Minimizing cybersickness in head-mounted display systems: causes and strategies review. In *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, pp. 154–163. IEEE, 2020.

[61] A. N. Ramaseri Chandra, F. El Jamiy, and H. Reza. A systematic survey on cybersickness in virtual environments. *Computers*, 11(4):51, 2022.

[62] J. T. Reason. Motion sickness adaptation: a neural mismatch model. *Journal of the royal society of medicine*, 71(11):819–829, 1978.

[63] J. T. Reason and J. J. Brand. *Motion sickness.* Academic press, 1975.

[64] L. Rebenitsch and C. Owen. Review on cybersickness in applications and visual displays. *Virtual Reality*, 20:101–125, 2016.

[65] G. M. Reger, J. S. McGee, C. van der Zaag, M. Thiebaux, J. G. Buckwalter, and A. A. Rizzo. A 3d virtual environment rod and frame test: the reliability and validity of four traditional scoring methods for older adults. *Journal of clinical and experimental neuropsychology*, 25(8):1169–1177, 2003.

[66] J. A. Richardson and T. E. Turner. Field dependence revisited i: Intelligence. *Educational psychology*, 20(3):255–270, 2000.

[67] O. N. Saracho. Cognitive style and kindergarten pupils' preferences for teachers. *Learning and Instruction*, 11(3):195–209, 2001.

[68] V. Sevinc and M. I. Berkman. Psychometric evaluation of simulator sickness questionnaire and its variants as a measure of cybersickness in consumer virtual environments. *Applied ergonomics*, 82:102958, 2020.

[69] K. Stanney, C. Fidopiastis, and L. Foster. Virtual reality is sexist: but it does not have to be. *Frontiers in Robotics and AI*, 7:4, 2020.

[70] K. M. Stanney, K. S. Kingdon, and R. S. Kennedy. Dropouts and aftereffects: Examining general accessibility to virtual environment technology. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 46, pp. 2114–2118. SAGE Publications Sage CA: Los Angeles, CA, 2002.

[71] R. C. Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.

[72] T. M. Therneau. *A Package for Survival Analysis in R*, 2023. R package version 3.5-7.

[73] T. Van Gemert, N. C. Nilsson, T. Hirzle, and J. Bergström. Sicknificant steps: A systematic review and meta-analysis of vr sickness in walking-based locomotion for virtual reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–36, 2024.

[74] R. A. A. Vingerhoets, M. De Vrijer, J. A. M. Van Gisbergen, and W. P. Medendorp. Fusion of visual and vestibular tilt cues in the perception of visual vertical. *Journal of neurophysiology*, 101(3):1321–33, Mar 2009. doi: 10.1152/jn.90725.2008

[75] E. Vittinghoff and C. E. McCulloch. Relaxing the rule of ten events per variable in logistic and cox regression. *American Journal of Epidemiology*, 165(6):710–718, mar 2007. doi: 10.1093/aje/kwk052

[76] G. Wang and A. Suh. User adaptation to cybersickness in virtual reality: a qualitative study. *Proceedings of the 27th European Conference on Information Systems*, pp. 8–14, 2019.

[77] S. Weech, C. M. Calderon, and M. Barnett-Cowan. Sensory down-weighting in visual-postural coupling is linked with lower cybersickness. *Frontiers in Virtual Reality*, 1:10, 2020.

[78] S. Weech, S. Kenny, and M. Barnett-Cowan. Presence and cybersickness in virtual reality are negatively related: a review. *Frontiers in psychology*, 10:415654, 2019.

[79] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[80] H. A. Witkin. Perception of body position and of the position of the visual field. *Psychological monographs: general and applied*, 63(7):i, 1949.

[81] H. A. Witkin and S. E. Asch. Studies in space orientation. iv. further experiments on perception of the upright with displaced visual fields. *Journal of Experimental Psychology*, 38(6):762, 1948.

[82] H. A. Witkin and D. R. Goodenough. Field dependence revisited. *ETS Research Bulletin Series*, 1977(2):i–53, 1977.

[83] H. A. Witkin, H. B. Lewis, M. Hertzman, K. Machover, P. B. Meissner, and S. Wapner. Personality through perception: An experimental and clinical study. 1954.

[84] O. P. K. R. E. . K. S. A. Witkin, H. A. *A manual for the embedded figures tests*. Consulting Psychologists Press, 1971.

[85] L. Yardley. Motion sickness susceptibility and the utilisation of visual and otolithic information for orientation. *European archives of oto-rhino-laryngology*, 247(5):300–4, 1990.

[86] S. D. Young, B. D. Adelstein, and S. R. Ellis. Demand characteristics of a questionnaire used to assess motion sickness in a virtual environment. In *Virtual Reality Conference, 2006*, pp. 97–102, 2006.

[87] L.-f. Zhang. Field-dependence/independence: cognitive style or perceptual ability?—-validating against thinking styles and academic achievement. *Personality and individual differences*, 37(6):1295–1311, 2004.