You Make Me Sick! The Effect of Stairs on Presence, Cybersickness, and Perception of Embodied Conversational Agents

Samuel Ang* Amanda Fernandez† Michael Rushforth‡ John Quarles§

University of Texas at San Antonio

Figure 1: View of the virtual environment with walls removed for view of the interior.

Abstract

Virtual reality (VR) technologies are used in a diverse range of applications. Many of these involve an embodied conversational agent (ECA), a virtual human who exchanges information with the user. Unfortunately, VR technologies remain inaccessible to many users due to the phenomenon of cybersickness: a collection of negative symptoms such as nausea and headache that can appear when immersed in a simulation. Many factors are believed to affect a user's level of cybersickness, but little is known regarding how these factors may influence a user's opinion of an ECA. In this study, we examined the effects of virtual stairs, a factor associated with increased levels of cybersickness. We recruited 39 participants to complete a simulated airport experience. This involved a simple navigation task followed by a brief conversation with a virtual airport customs agent in Spanish. Participants completed the experience twice, once walking across flat hallways, and once traversing a series of staircases. We collected self-reported ratings of cybersickness, presence, and perception of the ECA. We additionally collected physiological data on heart rate and galvanic skin response. Results indicate that the virtual staircases increased user level's of cybersickness and reduced their perceived realism of the ECA, but increased levels of presence.

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

*e-mail: samuel.ang@utsa.edu
†e-mail:amanda.fernandez@utsa.edu
‡e-mail:michael.rushforth@utsa.edu
§e-mail:john.quarles@utsa.edu

1 Introduction

Head mounted displays (HMD) are an increasingly available technology with an ever expanding domain of applications. These applications include in classroom learning environments [46], the treatment of psychological disorders [39, 43], remotely piloting drones [64], assistance in physical therapy [48] and language learning [19].

VR applications are often used in conjunction with embodied conversational agents (ECA). These agents are represented as humans within the virtual environment, and convey information through verbal, text-based, and animated feedback, often accepting responses from the user in return [47]. Opportunities to combine these technologies are plentiful given the many shared applications, such as education [24] and psychological assistance [37]. Researchers have found that an ECA's ability to convincingly display emotion and relate to users may impact the usability of the application [57]. For this reason, it is important to understand how various characteristics of the virual environment (e.g traversable geometry) may impact a user's relationship to ECAs in the scene.

Meanwhile, virtual reality applications are uniquely hindered by the phenomenon known as cybersickness, a term encompassing the variety of negative symptoms users can experience, such as nausea, dizziness, and eyestrain [34]. Little is known regarding how factors that exacerbate cybersickness may also impact a user's impression of an ECA. Cybersickness is believed to be negatively related with user levels of presence, the feeling of 'being there' [60]. If these factors can reduce the believably of the virtual environment, they could impact the believably of ECAs in particular, and by extension a user's ability to engage in productive and enjoyable interactions.

We designed a within-subjects user study to measure the impact of virtual staircase traversal on cybersickness, presence, and user interactions with an ECA. In our study, 39 participants completed a Spanish conversation task with an ECA in a simulated airport environment. This task was completed twice by each participant, once after navigating an empty virtual hallway, and once after navigating a hallway filled with staircases intended to evoke mild to moderate

2642-5254/23/$31.00 ©2023 IEEE
DOI 10.1109/VR55154.2023.00071

cybersickness. Rough terrain [3] and virtual stairs [20] have been demonstrated to produce worse cybersickness outcomes compared to flat surfaces in five minutes or less. Staircases in particular have also been shown to raise task difficulty [8].

During navigation, participants were asked to verbally rate their discomfort using the Fast Motion Sickness Scale [29] every thirty seconds, along with their level of presence using the Single Item Presence Scale (SIP) [11]. After the navigation task, participants completed a conversation in Spanish with a virtual airport customs agent. Post trial sickness levels were reported using the Simulator Sickness Questionnaire (SSQ) [28]. Presence ratings were collected using the Igroup presence questionnaire [49]. We additionally collected heard rate and galvanic skin response data as physiological metrics. User perception of the ECA was rated using the Conversational Agents Scale [59].

Results from the FMS indicate that participants felt greater levels of cybersickness while navigating the stairs condition compared to empty hallways. Participants also rated higher levels of presence in the staircase condition, but lower levels of perceived realism with regards to the ECA.

2 Background

2.1 Cybersickness

The term cybersickness refers to a series of uncomfortable symptoms that can manifest within virtual reality. Common symptoms include nausea, sweating, dizziness, and eyestrain [34]. These effects are more likely to arise when using a head-mounted display (HMD) than with common desktop monitors [52]. These symptoms can impact user task performance [5], enjoyment [51], and can force users to leave VR when intense enough. Thus, it is crucial to both uncover the underlying explanations for cybersickness, and develop techniques to mitigate the impacts. The sensory conflict theory, postural instability theory, and poison theory are three of the competing explanations for the existence of cybersickness [16]. The theory of sensory conflict theory attributes cybersickness to inconsistencies between visual sense input and the user's real-world movement. The postural instability theory instead argues that cybersickness stems from the human body's resistance to drastic changes in orientation. Finally, the poison theory attributes cybersickness to an evolutionary rejection of toxic substances that affected human perception.

Building upon these theories, researchers have developed various techniques and design principals for minimizing the impact of cybersickness. Some techniques aim to reduce the amount of unnecessary visual information available to the user by distorting or obscurring portions of the screen. This has been accomplished through digital filters that blur the screen [12, 35, 41, 45], and visual effects such as depth of field [14, 26]. Reducing field of view (FOV) by placing a filter around the user's peripheral vision has been utilized for this purpose as well [21, 27, 31]. Rest frames are another tool to combat cybersickness, objects in the simulation that remain stationary relative to the user's perspective [4, 13]. Many techniques share an inherent drawback of potentially obscuring important information. For example, these effects could prevent the user from reading text within the simulation. To reduce reliance on potentially problematic techniques, researchers have worked to develop guidelines for the design of comfortable virtual environments.

Researchers aim to reduce cybersickness with strong design principles regarding basic factors of the simulation. One such factor is the user locomotion method. Alternative locomotion methods such as teleportation [22, 36], and node-based travel [22] can ease the strain on users as they travel through the virtual environment. Virtual movement speed has also been demonstrated to impact a user's comfort in VR [33]. Other relevant factors include virtual object density [25], graphical realism [17, 44], and traversable geometry, such as stairs [20]. Like the reduction techniques previously discussed, design principles surrounding movement and environment layout may not be relevant to every simulation. By developing an understanding of these factors, researchers can better equip developers to create comfortable and inclusive simulations.

There are many approaches for measuring a user's level of cybersickness. These can come in the form of subjective questionnaires that the user completes after their VR experience, such as the Simulator Sickness Questionnaire [28]. This questionnaire features 16 subscales allowing users to rate individual symptoms, and can be used to calculate four overall scores describing nausea, oculomotor problems, disorientation, and total sickness. Some subjective metrics are simple enough that users can report values during the VR experience, such as with the Fast Motion Sickness Scale [30], which asks for a single value describing discomfort. In addition to subjective metrics, user cybersickness can be estimated using objective physiological readings. Heart rate, blink rate, and skin conductance have been found to positively correlate with ratings on the Simulator Sickness Questionnaire [32]. Physiological metrics such as stomach activity, and breathing have been used to predict reported reported cybersickness levels [18].

2.2 Embodied Conversational Agents

VR applications sometimes need to convey information through autonomous virtual agents that resemble humans. In addition to communicating through text, or auditory language, these virtual agents can make use of their body and facial features to perform gestures or communicate emotions. These entities are commonly referred to as Embodied Conversational Agents (ECA) [15]. ECAs have been used in a variety of applications, such as psychological evaluation [37], education [24], and practicing new languages [62]. As discussed by Baylor and Ebbers [6], ECAs carry possible advantages over real humans in educational settings, as they can be available on demand, and can be modified to appeal to particular groups.

Researchers have developed several methods of understanding a user's attitude towards an ECA. Adcock and Van Eck [1] developed the Attitude toward Tutoring Agent Scale (ATTS), from existing educational rating scales. Baylor and Ryu [7] proposed the Agent Persona Instrument, with questions meant to evaluate both informational usefulness and effective interaction. Both of these questionnaires are intended to assess pedagogical agents who assume the role of a teacher within the application. More recently, Wechsung et al. [59] have created the conversational agents scale (CAS). This scale was designed to be more inclusive and applicable to non-pedagogical agents; measuring likability, entertainment value, helpfulness, and naturalness. We used the CAS in our study to examine the impacts of virtual stairs on user perception of these four specific qualities.

Understanding how various factors may impact a user's perception of, and interactions with ECAs is important for many reasons. The perceived authenticity of an ECA's emotions may impact the overall usability of the application [57]. ECAs are employed to understand social behavior in controlled settings, such as understanding user anxiety in romantic encounters [42]. If external factors of the virutal environment affect user relationships to ECAs, then that information is crucial for designing such studies. In some cases, an ECA or collection of ECAs may be the central focus of the simulation, making them a key factor to maximizing the plausibility of the environment overall [9].

The increased immersion of HMDs could make ECAs seem more realistic, but HMDs also carry specific and unique challenges. As discussed in the previous section, using an HMD can increase cybersickness. Cybersickness is believed to impact other aspects of a user's perception, such as their feelings of presence with regards to the believably of the virtual environment [60]. Our study aims to expand knowledge on how traversing height altering geometry such as stairs may interact with other elements of perception towards an

ECA, as measured by the CAS.

2.3 Presence

Presence has been described as a user's feeling of 'being there' while immersed in a simulation, along with their suspension of disbelief, and level of involvement [53]. Higher levels of presence are generally desirable. Presence has been positively associated with enjoyment [56], and has been used as a metric to determine how closely a simulation mimics real-world experience [58].

Researchers have many available options when it comes to measuring presence. These include behavioral measures such as postural response [55], and subjective measures such as questionnaires. These questionnaires typically ask users to rate their feelings using a series of likert scales. Commonly used questionnaires include the Slater-Usoh-Steed Questionnaire [54], the Witmer-Singer Presence Questionnaire [63], and the more recent Igroup presence questionnaire (IPQ) [49] which was used in this study. In addition to subjective questionnaires, physiological metrics have been used in the past to measure presence from an objective standpoint. Changes in heart rate and skin conductance in particular have been found to be a reliable measure of presence within virtual environments [38] using the University College London (UCL) questionnaire and a study that modified user presence with changes in VR application framerate.

Many characteristics of a simulation are believed to influence presence, such as a user's virtual avatar [23], control interface [50], and visual realism [40]. However, we are not aware of any research examining how moving across different types of virtual geometry may affect a user's level of presence. These possible impacts are one focus of this study.

3 METHODS

Our within-subjects study features two conditions which we have dubbed the flat condition and the stairs condition. In the flat condition, users navigate empty hallways before arriving at a language speaking task. In the stairs condition, these hallways are populated by occasional staircases. Participants completed the simulation twice in counterbalanced order. Below we describe our methodology and resources in greater detail.

3.1 Virtual Environment

Each VR trial took place inside a simulated airport. Participants were informed that in the scenario, they had landed in Columbia to visit for two weeks. This scenario was selected as something simple to explain, and a situation in which even participants without a complete grasp of the Spanish language could see themselves needing to speak it. The first section is a series of hallways, with alternating right and left turns every 50 meters. These turns were introduced so that some level of engagement was required from the user to progress forward. These hallways were 5 meters wide and 5 meters tall. This study had two VR conditions, one in which the hallways were empty, another where they were populated by alternating ascending and descending stair cases every 5 meters. These staircases changed the user's elevation by 2.5 meters and were comprised of 20 steps each. We refer to these as the flat, and stairs conditions respectively. A user's view of the flat and stairs conditions is pictured in Figure 2

Movement through these hallways was controlled with touchpad input via the handheld HTC Vive controller. By placing a thumb to the edge of the touchpad, participants could determine their velocity relative to the direction they were facing. This control scheme was selected so that navigation was possible with a single hand, allowing for the other hand to be connected to our physiological sensors. Movement speed was set to a 4.25 m/s. We selected this speed based off of the HMD game The Talos Principle VR. This game classifies a movement speed of 4.25 m/s as "slow" and has been previously used to study the effects of various walking speeds in VR [2].

Figure 2: Participant's view of hallway for each condition. Flat hallways are pictured on the left while hallways with stairs are pictured on the right.

The second section of our simulation featured a virtual customs agent standing behind a desk. This customs agent was capable of speaking pre-recorded voice lines and lip syncing to the audio. The agent was additionally equipped with a simple idle animation that had them occasionally turning their head and swaying back and forth. This area of the simulation is pictured in Figure 3

Figure 3: Participant's view of the virtual customs agent.

3.2 Task

Participants completed two tasks during each VR session. The first was a navigation task through the airport hallways. These hallways were either empty, or obstructed by occasional ascending or descending staircases depending on the condition of the trial. Both variants of these hallways are depicted in Figure 2. Participants navigated these hallways for four minutes. From a previous study with a similar virtual navigation task, we found that nearly half of participants who needed to end the simulation early did so shortly after 4 minutes. This duration was selected in order to maximize the number of participants who would be able to complete the following language task and produce necessary data.

After four minutes of navigation, participants were shown a brief fade to black through the HMD as if they had closed their eyes, before being teleported several meters from the customs agent. Participants were then asked to approach the desk using the same movement metaphor they used before. Once close enough, the customs agent would ask for the participant's passport. A passport would then appear in the participant's hand, and could be placed on the desk by extending their arm forward. This was not necessary to proceed with the conversation, however. After registering a user's verbal response with the HMD's built in microphone, the agent proceeded to the next question. This lasted until all dialogue was exhausted. The agent's questions are listed below in the order they were asked and answered.

- Buenos días. ¿Su pasaporte, por favor? (Hello. Your passport please?)
- ¿Cuál es el propósito de su visita? (What is the purpose of your visit?)
- ¿Cuánto tiempo piensa estar aquí? (How long do you plan to be here?)
- Bueno. ¿Dónde se quedará? (Good. Where will you stay?)
- ¿Eres un miembro de la Fuerza Aérea de los Estados Unidos? (Are you a member of the United States Air Force?)
- Mencionó que viajará y regresará a Colombia por unos días. ¿Dónde se va a quedará en ese momento? (You mentioned that you will travel and return to Colombia for a few days. Where will you stay at that time?)

After all questions were answered, the agent said the following to conclude the language task.

- Ah sí, ya veo. Completó ambas direcciones en el formulario. Muy bien, disfruta su viaja. Esa es la salida. (Ah yes, I see. You filled out both addresses on the form. Very good, enjoy your trip. That is the exit.)

Participants were then asked to remove the HMD.

3.3 Procedure

Participants were given an explanation of the task and the control scheme for navigating the environment. They were also informed that they could ask to end the study at any point if they become too uncomfortable. We had zero participants drop out at any point in the study. Next, they were fitted with the HTC Vive Pro headset, physiological monitoring equipment, and a handheld controller.

Two training simulations were conducted prior to beginning the experiment. The first was a minute long period for participants practice with the movement controls in the flat terrain condition. The second was a portion of the dialogue exchange with the ECA while the HMD display was switched off. The first two questions were asked and answered in Spanish to test the HMD's built in microphone and demonstrate the speed at which the ECA was capable of responding. Participants were kept from seeing the ECA so as to minimize the impact of their impression during training.

After training was complete, the equipment was removed and the participant filled out a background questionnaire and an initial SSQ to serve as a baseline. Values from the baseline SSQ were subtracted from final SSQ scores before the final analysis. Once complete, the HTC Vive was re-equipped and the simulation was initiated with either flat terrain or stairs depending on counterbalanced order. Using the handheld controller, participants navigated the airport hallways for a period of 4 minutes. Every 30 seconds, the study administrator asked for a verbal rating of discomfort on the Fast Motion Sickness Scale (FMS) from 0 to 10, and a rating of presence using the Single Item Presence Scale (SIP) from 0 to 10.

Once complete, participants removed their equipment and completed a series of post exposure questionnaires. This was comprised of the Simulator Sickness Questionaire (SSQ), Conversational Agents Scale (CAS), and Igroup Presence Questionnaire (IPQ). Once the questionnaires were completed, participants were given a five minute break. This break was extended by one minute at a time until participants verbally reported a score of 0 on the FMS scale. The simulation was then repeated with the next hallway type. A simplified outline of the study procedure is illustrated in Figure 4

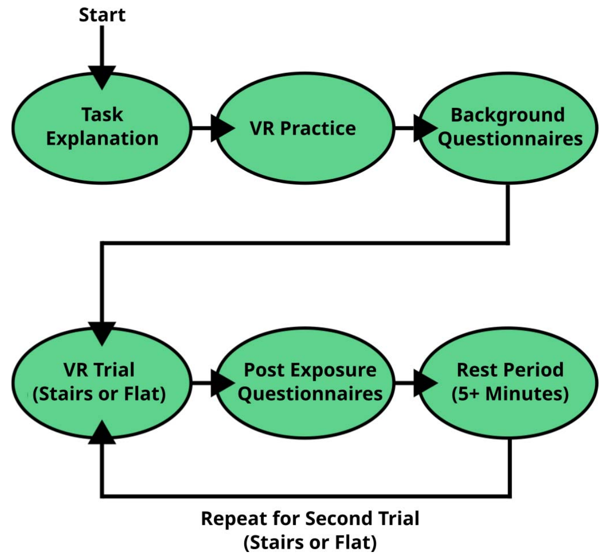

Figure 4: Outline of study procedure.

3.4 Hypotheses

We made three hypotheses regarding the outcome of our study.

- H1: Users will report greater levels of cybersickness for the stairs condition than for the flat condition.
- H2: Users will report lower levels of presence for the stairs condition than for the flat condition.
- H3: Staircases will negatively impact user ratings of ECA naturalness.

H1 was made given the existing work that has identified height altering terrain [3] as a contributing factor to cybersickness. Staircases may reduce user postural stability, and introduce an additional dimension of movement that does not match the users physical experience, both of which are theorized to be responsible for causing cybersickness [16]. H2 follows from H1, and was made because presence tends to be negatively related with cybersickness [60]. Finally, H3 was made because realism is an element of presence. If the realism of the environment overall is affected, then it would stand to reason that the realism of the ECA in particular is affected as well.

3.5 Apparatus

The simulated airport was created within the Unity game engine. The HTC Vive Pro Eye was used as the HMD for our study. Virtual locomotion was accomplished through use of a handheld HTC controller. Heart rate and GSR were collected using the Neulog NUL-208 and NU-217 sensors. An HP Omen laptop running Windows 10 was used to run the simulation. This laptop came equipped with an AMD Ryzen 7 4800H processor, an NVIDIA 1660ti graphics card, and 16 GB of RAM. Our simulation was built so that the HMD could maintain a constant framerate of 90hz in conjunction with the latpop's hardware.

4 Pilot Study

An initial pilot study with five participants was run to reveal potential issues. This lead us to make three adjustments to the design of our study. Originally, the virtual environment was designed such that users seamlessly arrived at the customs agent's desk after a fixed time period. This was done by dynamically changing the length of the virtual hallway once participants neared their time limit. This method was incapable of standardizing time spent in navigation with a reasonable threshold. Participants consistently completed the navigation too late or too early. For the final study, we teleported users to their destination once the four minutes were up after a brief fade to black. To standardize the transition further, we made the hallway long enough so that no participant could reach their destination on time and would necessarily experience the teleportation effect.

The first question from the ECA asks participants to hand over their passport, at which point a passport appears in the participant's right hand. Passing this to the agent was originally required to begin the interaction, however all pilot study participants forgot to complete this step, so it was made optional for the full study. Participants were still informed that a passport would appear in their had but that passing it across the desk was not necessary. The conversation is instead initiated by the participant's first verbal response into the HMD microphone.

Finally, one participant ran into an issue during one of their trials where the customs agent failed to recognize their speech, and remained silent. To ensure this did not happen in the future, we included an emergency button that the study administer could use to manually activate the agent's next line in the dialogue sequence. This ensured that the dialogue task proceeded smoothly without the participant becoming aware of software failure.

4.1 Participants

We recruited 39 (14 M, 25 F) participants for this study. These were recruited from the local area and through Spanish organizations at our university. The mean age of participants was 23.95 with a standard deviation of 10.10 years. Twelve of these participants identified as native speakers of Spanish. Of those who were not native speakers had taken an average of 2.79 Spanish classes with a standard deviation of 2.74. Participants were compensated $35 an hour for their time.

5 Metrics

In this section we discuss the various questionnaires we employed to understand user levels of cybersickness, presence, and attitudes towards the embodied conversational agent.

5.1 Cybersickness Metrics

During each four minute navigation task, participants were asked to verbally rate their level of sickness using the Fast Motion Sickness (FMS) Scale [30]. This was rated on scale of 0 (none) to 10 (severe), and asked every thirty seconds. Heart rate and GSR were collected automatically using a Neulog NUL-208 and NU-217, which connected to the user's free hand via a series of wires.

After leaving VR, participants rated their symptoms using the Simulator Sickness Questionnaire (SSQ) [28]. The SSQ is comprised of 16 scales rating symptoms from 0 (none) to 3 (severe). Responses to these questions can be used to generate sub scores describing nausea, oculomotor issues, disorientation, and total sickness.

5.2 Presence Metrics

While immersed in VR, participants were asked to verbally rate their level of presence from 0 (low) to 10 (high) using the Single Item Presence Scale (SIP) [11]. This was done every thirty seconds immediately following their FMS rating. Participants were asked "On a scale of 0 – 10 where 0 is 'not at all present' and 10 is 'totally present', to what extent did you feel present in the environment, as if you were really there?"

After completing the VR session, participants rated their levels of presence using the Igroup presence questionnaire [49]. This questionnaire contains 14 ratings from 0 (low) to 6 (high) describing spatial presence, involvement, realism, and general presence.

5.3 ECA Metrics

User reports on the interaction with the ECA was collected with the Conversational Agents Scale (CAS) [59]. This questionnaire is made up of 25 seven-point responses. Each response has users grade the agent based on two descriptions. For example, "the agent was pleasant vs unpleasant." These scores make up six dimensions of likeability, entertainment, helpfulness, naturalness, trust, and perceived task difficulty.

6 Results

In this section we report our findings from each metric discussed in the previous section. As normality could not be assumed, we used a Wilcoxon signed-rank test to determine statistical significance.

6.1 FMS Results

Each participant spent 4 minutes completing their navigation task in VR. Every 30 seconds, they were asked to verbally rate their level of sickness from 0 (no sickness) to 10 (severe sickness) using the FMS [30]. Over the course of the navigation task, participants reported eight scores. We created a single overall FMS score for each participant by averaging their eight ratings. Mean overall FMS scores were higher for the stairs condition (M=1.77, SD = 1.75) compared to those from the flat condition (M = 1.09, SD = 1.35). In our analysis we performed a Wilcoxon signed-rank test and found these results to be statistically significant (V = 86.5, p =.003, Cohen's d = .434), allowing us to reject the null hypothesis.

For each participant, we also collected the maximum FMS score they reported over the course of their navigation experience. Mean maximum scores were once again higher for the stairs condition (M=2.60, SD = 2.30) than the flat condition (M=1.59, SD = 1.71). A Wilcoxon signed-rank test again found these results to be statistically significant (V=91.5, p = .006, Cohen's d = .493), allowing us to reject the null hypothesis.

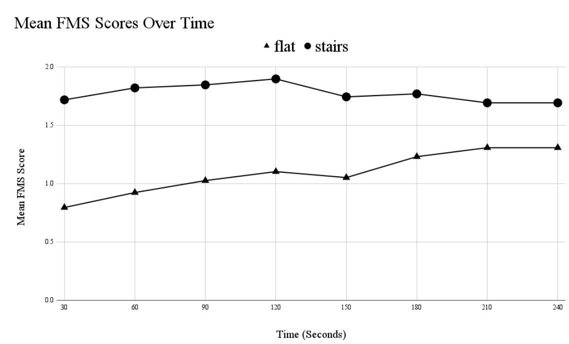

Figure 5: User FMS scores collected over time at 30 second intervals.

6.2 SSQ Results

After each VR trial, participants completed a simulator sickness questionnaire. SSQ nausea, oculomotor, disorientation, and total scores were virtually identical between each condition. After performing a Wilcoxon signed-rank test on each of the subscore and total score pairings, we did not find the differences to be statistically significant between nausea (V = 139.5, p =.976, Cohen's d = .097) oculomotor (V = 174.5, p=.522, Cohen's d = .095) disorientation (V = 97.5, p=.081, Cohen's d = .240) or total scores (V = 201, p=.3618, Cohen's d = .153). The mean scores for each category are displayed in Table 1

We additionally compared the 16 individual ratings from the SSQ against one another and found one that dizziness with eyes closed

	Flat	Stairs
Nausea	3.92±.79	4.02±.81
Ocularmotor	5.42±1.00	5.43±1.10
Disorientation	5.25±.94	5.30±.90
Total	3.56±1.25	5.07±.90

Table 1: Mean SSQ subscores for each condition.

was higher for the stairs condition (M=.23,SD=.48) than the flat condition (M=.05,SD=.22). A Wilcoxon signed-rank test found this to be statistically significant (V = 0, p = .026, Cohen's d = .475).

6.3 SIP Results

In addition to an FMS rating, participants were asked to provide a rating of presence using the SIP scale [11] from 0 (low) to 10 (high). These scores were asked every 30 seconds, resulting in 8 scores for each participant per condition. We calculated the mean of these 8 scores in order to reach an overall presence score describing the participants' experience for each condition. Mean overall presence scores from SIP were higher for the stairs condition (M= 5.15, SD = 2.19) than for the flat condition (M=4.79, SD = 2.28).

A Wilcoxon signed-rank test found these results to be statistically significant (V = 125, p = .010, Cohen's d = .161) allowing us to reject the null hypothesis.

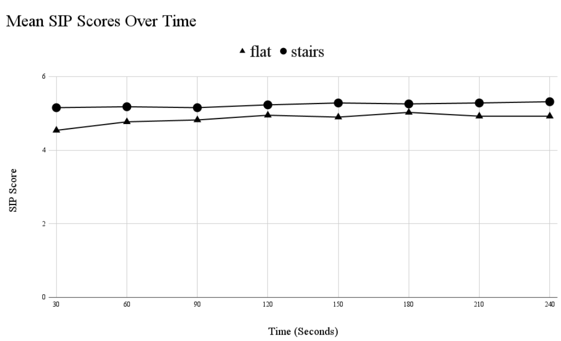

Figure 6: User SIP scores collected over time at 30 second intervals.

6.4 IPQ Results

Results from the IPQ support those we collected from the SIP. Spatial presence subscores were higher on average when reported after completing the stairs condition (M=2.70, SD=.78) than the flat condition (M=2.45. SD=.78). A Wilcoxon signed-rank test found this difference to be statistically significant (V = 153, p=.023, Cohen's d = .315).

We did not detect any significant differences between involvement subscores between the flat (M=2.33, SD = .962) and stairs conditions (M=2.26, SD=.942) via Wilcoxon signed-rank test (V = 279.5, p=.540, Cohen's d = .074) This was also true for realism subscores between the flat (M=1.81, SD = .623) and stairs conditions (M=1.94, SD=.716) via Wilcoxon signed-rank test (V = 139.5, p = .148, Cohen's d = .181). Reports for the general presence subscore after the stairs condition were once again higher (M=2.02, SD = .903) compared to those following the flat condition (M=1.82, SD = 1.07), but were not significant (V = 95.5, p = .165, Cohen's d = .206).

6.5 Physiological Results

For the duration of each trial's navigation task, we collected heart rate (HR) and galvanic skin response (GSR) data. Both these values were collected at a rate of 10hz over four minutes, yielding 4800 readings for each participant, and 2400 for each metric. For each participant, we averaged these 2400 values to get an overall HR and GSR value. These values were used in our final analysis.

Mean HR readings were slightly higher for participants during the flat condition (M=75.03, SD=12.90) compared to the stairs condition (M=72.33, SD=11.13). A Wilcoxon signed-rank test found this difference to be statistically insignificant (V = 507, p = .105, Cohen's d = .224).

Mean GSR readings were higher for the stairs condition than the flat condition (M=2.68, SD=2.29) compared to the stairs condition (M=3.01, SD=2.19). This results were not statistically significant (V = 251, p = .052, Cohen's d = .146), though only barely.

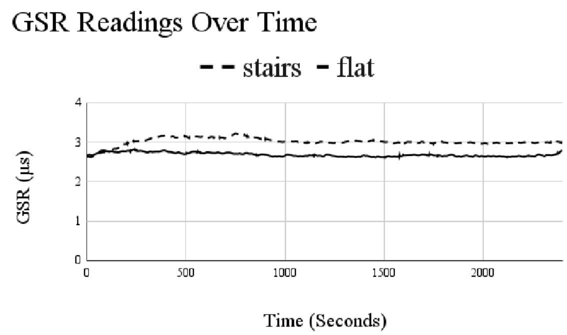

Figure 7: User GSR readings collected over time at 10hz.

6.6 CAS Results

CAS subscores were mostly identical between the two conditions across likability, entertainment, helpfulness, naturalness, trust, and perceived task difficulty subscores. Exact scores are listed in Table 2. No differences in user scores were determined to be statistically significant for entertainment (V = 274, p = .507, Cohen's d = .006), likability (V = 208, p = .622, Cohen's d = .027), helpfulness (V = 231, p = .259, Cohen's d = .050), naturalness (V = 389, p = .120, Cohen's d = .151), trust (V = 268.5, p = .837, Cohen's d = .055) or task difficulty (V = 379, p = .475, Cohen's d = .107). Individual items from the CAS were individually compared in the same way.

However, we found one statistically significant rating. When asked to judge the agent on realism on a scale of 1 (unrealistic) to 7 (realistic), participants provided higher ratings on average after the flat condition (M=4.41, SD=1.39) than after the stairs condition (M=3.97, SD=1.53). A Wilcoxon signed-rank test determined the differences in these ratings to be statistically significant (V = 133, p = .006, Cohen's d = .298). In accounting for order effects, we compared participants' first and second responses to this question regardless of which condition they completed first. A Shapiro-Wilk test found that neither first nor second responses were normally distributed. We performed a Wilcoxon signed-rank test and could not find evidence of a significant order effect (V = 65.5, p = 0.609, Cohen's d = .052).

To further explore the differences between the perceived realism of the ECA, we calculated the Pearson's correlation coefficient between user realism scores from the CAS and several other metrics. First we examined the relationship between CAS realism and user

	Flat	Stairs
Entertainment	3.92±0.79	4.02±0.81
Likability	5.42±1.00	5.43±1.10
Helpfullness	5.25±0.94	5.30±0.90
Naturalness	3.56±1.25	5.07±0.900
Trust	5.02±0.93	5.07±0.900
Difficulty	4.51±1.32	4.38±1.33

Table 2: Mean Conversational Agent Scale subscores for each condition.

FMS scores. For results from the flat condition, we did not find a significant correlation between CAS realism and average (r(37) = -.056, p=.734), final r(37) = -.067, p=.685), or maximum (r(37) = -.049, p=.767) participant FMS scores. Negative correlations between CAS realism and FMS were more exaggerated for the stairs condition. Only the correlation between CAS realism and final FMS scores were significant (r(37) = -.361, p=.024). Correlations with average (r(37) = -.293, p=.070) and maximum (r(37)=-.235, p = .159) FMS scores were not significant.

We also examined the relationships between CAS realism and each subscore from the IPQ. For scores from the flat condition, we found three significant correlations between CAS realism and IPQ spatial presence (r(37) = .538, $p < .001$), realism (r(37) = .416 p=.008) and general presence r(37) = 0.509 $p < .001$). We found similar results from the stairs condition, with significant correlations between CAS realism and IPQ realism (r(37) = .461, p=.003) and general presence (r(37) = .381, p=.017). The correlation between CAS realism and IPQ spatial presence (r(37) = .246, p = .132) was not significant for the stairs condition however.

7 Discussion

7.1 FMS Discussion

Results from the FMS support our **H1**, that the introduction of stairs into the virtual scene would increase user cybersickness. These results match findings from previous studies [3, 20], bolstering the notion that interruptions in simulated verticality can impact user comfort. Two explanatory theories of cybersickness could explain this outcome. These results are in line with the sensory conflict theory given that vertical displacement introduces an extra axis of motion the user is not traveling along in physical space. The postural instability theory is also relevant as changes in simulated position may reduce a user's balance. The overall effect sizes on mean (Cohen's d=.434) and maximum (Cohen's d = .493) were small. This is in line with our objective of minimizing the number of participant dropouts during the navigation task to maximize the number of interactions with the ECA.

Mean FMS scores appear to converge over time, as illustrated in Figure 5. Every thirty seconds, the distance between mean scores decreased with the exception of the distances between the final two pairs of scores which were identical. Given additional time in VR, it is unknown if these scores would eventually converge completely or remain in parallel. Many participants expressed verbal surprise after climbing a staircase for the first time. This initial shock may explain the initial gap between the scores of each condition, but more research is needed to confirm this quantitatively.

7.2 SSQ Discussion

Mean SSQ subscores failed to exceed a value of 10 for both conditions and can be interpreted as negligible to minimal [10]. Traversable geometry did not appear to greatly affect SSQ scores. Results from a Wilcoxon signed-rank test comparing these subscores were not statistically significant, and effect size of terrain was negligible for each. These results were surprising given the clear difference in FMS ratings between the two conditions. Our study design may be responsible for the lack of findings from SSQ ratings. After completing their navigation, participants completed a conversational task that did not involve moving through the virtual world. Over the course of this conversation task, symptoms of cybersickness could have subsided. We did not ask users for FMS ratings during the conversation task so as not to interrupt their experience, so we do not know for certain if these scores declined during that period of time.

7.3 Presence Discussion

We detected a statistically significant effect of terrain type on user presence using the Single Item Presence Scale [11] and the IPQ. Interestingly, presence scores for the stairs condition were significantly higher than those from the flat condition, contradicting **H2**. This is unexpected given that presence and cybersickness tend to be negatively related [60]. Higher levels of presence have also been associated with higher levels of enjoyment however [56]. It may be the case that the virtual stairs proved more exciting, enjoyable, or stimulating to users as they completed their navigation task. Further research would be needed to uncover the underlying cause behind this difference. Reported presence did not shift extensively over time. Twenty-four of our thirty-nine participants did not report a score more than one point away from their initial report for either trial. The effect on presence detected through SIP scores was extremely small (Cohen's d = .161). However, this effect was slightly higher when looking at spatial presence scores from the IPQ (Cohen's d = .315). Exaggerated forms of virtual terrain could exacerbate this difference, but future research is necessary before making such conclusions.

7.4 CAS Discussion

We did not detect any statistically significant differences in CAS subscores between the flat and stairs conditions. Because of this we were unable to confirm **H3**. This lead us to examine each individual rating on the CAS for possible changes in the user perspective. The only statistically significant difference we found was that of perceived realism of the ECA, which was higher when reported after the flat condition. We additionally found a significant negative correlation between CAS realism and final FMS scores, but only for scores from the stairs condition. A future study focused on the direct impacts of cybersickness on ECA perception may be able to determine the existence of a causal link, but that can not be concluded from correlation alone.

CAS realism was positively correlated with general presence, and realism subscores from the IPQ for both conditions. CAS realism was also positively correlated with the spatial presence for the flat condition in particular. It appears that there is a positive association between user impressions of realism overall, and ECA realism in particular. It is unusual that the stairs condition seemed to increase user presence while decreasing perceived ECA realism given the observed positive correlation. One explanation is that this correlation may have more to do with how users perceive the questions than the actual effect of the condition however.

8 Limitations

This study is limited by several elements stemming from its design to sample population. As discussed prior, we designed our study with the intention of inducing only a management amount of cybersickness so that participants would feel well enough to complete the conversation task. As a result, participants only engaged in navigation tasks for a brief period of four minutes at a time. When using VR for other purposes, users may need to remain immersed for much longer. Our results may not accurately describe longer periods in VR. As a result of our design, overall sickness levels were low, and we had no dropouts over the course of our entire study. In future work we plan to examine the impact of more intense cybersickness in riskier environments.

We employed the Conversational Agents Scale (CAS) [59] so that participants could quantitatively score their perception of the ECA accross several categories. This scale was described as a "work in progress" by the original authors and may not be ideal for our purposes. This questionnaire was selected because the others we found were directed more specifically at pedagogical agents [1, 7]. We wish to highlight the need for robust and standardized questionnaires for collection of subjective data regarding the perception of ECAs. This questionnaire is also originally written in German, though it has been employed for English speaking participants in the past [61].

The conversation task was completed in Spanish. The majority of our participants were not native speakers, and were using this scenario as a way to practice a language they were learning. Interacting with an ECA without a full grasp of the language may impact perception in unpredictable ways.

Finally, the study relies on a small sample size with just 39 participants. Most of these participants were recruited through Spanish organizations in the university, resulting in a low mean age. The gender balance of our sample size was also not even with 25 women and 14 men participating.

9 Observations

In this section we discuss spare observations and notes taken during each trial. We did not instruct participants to speak during the VR trial unless it was for the purposes of reporting a score or speaking to the ECA. Participants sometimes made comments of their own volition however. Many participants expressed excitement while traversing virtual staircases. This was particularly common when participants arrived at a descending staircase. Several participants also remarked that their sickness levels seemed to increase rapidly each time they descended a staircase. One participant remarked that he felt the staircases reduced his level of eyestrain by influencing the natural direction of his gaze. The possibility of manipulating a user's gaze via changes in height altering geometry may be interesting to explore in future work, though it is unclear what applications this would have.

10 Conclusion

Our study examined the effects of virtual staircases on cybersickness, presence, and user perception of embodied conversational agents. Participants completed an airport Spanish dialogue task after navigating a series of either empty hallways, or hallways populated with staircases. Results from subjective metrics indicate that participants felt significantly greater levels of cybersickness while navigating the stairs condition. However, results from the post-exposure questionnaires indicate that the stairs condition increased user presence ratings. These results may be explained by greater levels of postural instability, or greater levels of sensory conflict induced by the presence of virtual stairs. User scores regarding the ECA itself were largely inconclusive. On an individual scale, participants rated level ECA realism significantly higher after navigating flat hallways, compared to the stairs condition. With these results in mind, we recommend that developers remain aware that height altering geometry can increase cybersickness, and may negatively impact the realism of embodied conversational agents in the scene. We advise that they also keep in mind that altering geometry may also provide benefits in terms of user presence, though more research is needed to understand why.

Acknowledgments

This work was funded through grants from the National Science Foundation (IIS 2007041, IIS 2211785). This organization had no input regarding the planning or details of the study.

References

[1] A. B. Adcock and R. N. Van Eck. Reliability and factor structure of the attitude toward tutoring agent scale (attas). *Journal of Interactive Learning Research*, 16(2):195–217, 2005.

[2] A. Agić, E. Murseli, L. Mandić, and L. Kapov. The impact of different navigation speeds on cybersickness and stress level in vr. *Journal of Graphic Engineering and Design*, 11(1):5, 2020.

[3] S. Ang and J. Quarles. You're in for a bumpy ride! uneven terrain increases cybersickness while navigating with head mounted displays. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 428–435. IEEE, 2022.

[4] P. Bala, D. Dionísio, V. Nisi, and N. Nunes. Visually induced motion sickness in 360° videos: Comparing and combining visual optimization techniques. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 244–249. IEEE, 2018.

[5] E. M. Barhorst-Cates, K. M. Rand, and S. H. Creem-Regehr. The effects of restricted peripheral field-of-view on spatial learning while navigating. *PloS one*, 11(10):e0163785, 2016.

[6] A. Baylor and S. Ebbers. Evidence that multiple agents facilitate greater learning. *Artificial intelligence in education: Shaping the future of learning through intelligent technologies*, pages 377–379, 2003.

[7] A. Baylor and J. Ryu. The api (agent persona instrument) for assessing pedagogical agent persona. In *EdMedia+ innovate learning*, pages 448–451. Association for the Advancement of Computing in Education (AACE), 2003.

[8] L. Berger and K. Wolf. Wim: fast locomotion in virtual reality with spatial orientation gain & without motion sickness. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, pages 19–24, 2018.

[9] I. Bergström, S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater. The plausibility of a string quartet performance in virtual reality. *IEEE transactions on visualization and computer graphics*, 23(4):1352–1359, 2017.

[10] 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)*, pages 464–467. IEEE, 2020.

[11] S. Bouchard, G. Robillard, J. St-Jacques, S. Dumoulin, M.-J. Patry, and P. Renaud. Reliability and validity of a single-item measure of presence in vr. In *The 3rd IEEE international workshop on haptic, audio and visual environments and their applications*, pages 59–61. IEEE, 2004.

[12] P. Budhiraja, M. R. Miller, A. K. Modi, and D. Forsyth. Rotation blurring: use of artificial blurring to reduce cybersickness in virtual reality first person shooters. *arXiv preprint arXiv:1710.02599*, 2017.

[13] Z. Cao, J. Jerald, and R. Kopper. Visually-induced motion sickness reduction via static and dynamic rest frames. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 105–112. IEEE, 2018.

[14] K. Carnegie. Mitigating visual discomfort on head mounted displays using estimated gaze dependent depth of field. 2015.

[15] J. Cassell. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine*, 22(4):67–67, 2001.

[16] S. Davis, K. Nesbitt, and E. Nalivaiko. A systematic review of cybersickness. In *Proceedings of the 2014 conference on interactive entertainment*, pages 1–9, 2014.

[17] S. Davis, K. Nesbitt, and E. Nalivaiko. Comparing the onset of cybersickness using the oculus rift and two virtual roller coasters. In *Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015)*, volume 27, page 30, 2015.

[18] M. S. Dennison, A. Z. Wisti, and M. D'Zmura. Use of physiological signals to predict cybersickness. *Displays*, 44:42–52, 2016.

[19] T. K. Dhimolea, R. Kaplan-Rakowski, and L. Lin. A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, pages 1–15, 2022.
[20] J. L. Dorado and P. A. Figueroa. Ramps are better than stairs to reduce cybersickness in applications based on a hmd and a gamepad. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 47–50. IEEE, 2014.
[21] A. S. Fernandes and S. K. Feiner. Combating vr sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 201–210. IEEE, 2016.
[22] M. J. Habgood, D. Moore, D. Wilson, and S. Alapont. Rapid, continuous movement between nodes as an accessible virtual reality locomotion technique. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 371–378. IEEE, 2018.
[23] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social vr. In *2017 IEEE symposium on 3D user interfaces (3DUI)*, pages 233–234. IEEE, 2017.
[24] S. Hobert and R. Meyer von Wolff. Say hello to your new automated tutor–a structured literature review on pedagogical conversational agents. 2019.
[25] Q. C. Ihemedu-Steinke, S. Rangelova, M. Weber, R. Erbach, G. Meixner, and N. Marsden. Simulation sickness related to virtual reality driving simulation. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 521–532. Springer, 2017.
[26] C. James and L. E. Potter. The effects of post-processing techniques on simulator sickness in virtual reality.
[27] N. Kala, K. Lim, K. Won, J. Lee, T. Lee, S. Kim, and W. Choe. P-218: An approach to reduce vr sickness by content based field of view processing. In *SID Symposium Digest of Technical Papers*, volume 48, pages 1645–1648. Wiley Online Library, 2017.
[28] 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.
[29] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Human factors*, 53(4):415–426, 2011.
[30] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Human factors*, 53(4):415–426, 2011.
[31] N.-G. Kim and B.-S. Kim. The effect of retinal eccentricity on visually induced motion sickness and postural control. *Applied Sciences*, 9(9):1919, 2019.
[32] Y. Y. Kim, H. J. Kim, E. N. Kim, H. D. Ko, and H. T. Kim. Characteristic changes in the physiological components of cybersickness. *Psychophysiology*, 42(5):616–625, 2005.
[33] K. K. Kwok, A. K. Ng, and H. Y. Lau. Effect of navigation speed and vr devices on cybersickness. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 91–92. IEEE, 2018.
[34] J. J. LaViola Jr. A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin*, 32(1):47–56, 2000.
[35] Y.-X. Lin, R. Venkatakrishnan, R. Venkatakrishnan, E. Ebrahimi, W.-C. Lin, and S. V. Babu. How the presence and size of static peripheral blur affects cybersickness in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 17(4):1–18, 2020.
[36] G. Loup and E. Loup-Escande. Effects of travel modes on performances and user comfort: a comparison between armswinger and teleporting. *International Journal of Human–Computer Interaction*, 35(14):1270–1278, 2019.
[37] J. Martínez-Miranda, A. Martínez, R. Ramos, H. Aguilar, L. Jiménez, H. Arias, G. Rosales, and E. Valencia. Assessment of users' acceptability of a mobile-based embodied conversational agent for the prevention and detection of suicidal behaviour. *Journal of medical systems*, 43(8):1–18, 2019.
[38] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks Jr. Physiological measures of presence in stressful virtual environments. *Acm transactions on graphics (tog)*, 21(3):645–652, 2002.
[39] B.-A. J. Menelas, C. Haidon, A. Ecrepont, and B. Girard. Use of virtual reality technologies as an action-cue exposure therapy for truck drivers suffering from post-traumatic stress disorder. *Entertainment computing*, 24:1–9, 2018.
[40] M. Newman, B. Gatersleben, K. Wyles, and E. Ratcliffe. The use of virtual reality in environment experiences and the importance of realism. *Journal of environmental psychology*, 79:101733, 2022.
[41] G.-Y. Nie, H. B.-L. Duh, Y. Liu, and Y. Wang. Analysis on mitigation of visually induced motion sickness by applying dynamical blurring on a user's retina. *IEEE transactions on visualization and computer graphics*, 2019.
[42] X. Pan, M. Gillies, C. Barker, D. M. Clark, and M. Slater. Socially anxious and confident men interact with a forward virtual woman: an experimental study. *PloS one*, 7(4):e32931, 2012.
[43] H. M. Peperkorn, J. E. Diemer, G. W. Alpers, and A. Mühlberger. Representation of patients' hand modulates fear reactions of patients with spider phobia in virtual reality. *Frontiers in psychology*, 7:268, 2016.
[44] M. Pouke, A. Tiiro, S. M. LaValle, and T. Ojala. Effects of visual realism and moving detail on cybersickness. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 665–666. IEEE, 2018.
[45] W. Qionghua, W. Hui, and W. Qiang. Some experimental results of relieving discomfort in virtual reality by disturbing feedback loop in human brain. *arXiv preprint arXiv:1903.12617*, 2019.
[46] A. B. Ray and S. Deb. Smartphone based virtual reality systems in classroom teaching—a study on the effects of learning outcome. In *2016 IEEE eighth international conference on technology for education (T4E)*, pages 68–71. IEEE, 2016.
[47] Z. Ruttkay, C. Dormann, and H. Noot. Embodied conversational agents on a common ground: A framework for design and evaluation. In *From brows to trust: evaluating embodied conversational agents*, pages 27–66. Kluwer, 2004.
[48] D. Saldana, M. Neureither, A. Schmiesing, E. Jahng, L. Kysh, S. C. Roll, and S.-L. Liew. Applications of head-mounted displays for virtual reality in adult physical rehabilitation: a scoping review. *The American Journal of Occupational Therapy*, 74(5):7405205060p1–7405205060p15, 2020.
[49] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001.
[50] J. Seibert and D. M. Shafer. Control mapping in virtual reality: effects on spatial presence and controller naturalness. *Virtual Reality*, 22(1):79–88, 2018.
[51] D. M. Shafer, C. P. Carbonara, and M. F. Korpi. Factors affecting enjoyment of virtual reality games: a comparison involving consumer-grade virtual reality technology. *Games for Health Journal*, 8(1):15–23, 2019.
[52] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson. Virtual reality induced symptoms and effects (vrise): Comparison of head mounted display (hmd), desktop and projection display systems. *Displays*, 29(2):58–69, 2008.
[53] M. Slater and M. Usoh. Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, 2(3):221–233, 1993.
[54] M. Slater, M. Usoh, and A. Steed. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 3(2):130–144, 1994.
[55] R. H. So and G. K. Chung. Sensory motor responses in virtual environments: Studying the effects of image latencies for target-directed hand movement. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 5006–5008. IEEE, 2006.
[56] S. Sylaiou, K. Mania, A. Karoulis, and M. White. Exploring the relationship between presence and enjoyment in a virtual museum. *International journal of human-computer studies*, 68(5):243–253, 2010.
[57] S. ter Stal, L. L. Kramer, M. Tabak, H. op den Akker, and H. Hermens. Design features of embodied conversational agents in ehealth: a literature review. *International Journal of Human-Computer Studies*, 138:102409, 2020.
[58] A. Wagler and M. D. Hanus. Comparing virtual reality tourism to real-life experience: Effects of presence and engagement on attitude and enjoyment. *Communication Research Reports*, 35(5):456–464, 2018.
[59] I. Wechsung, B. Weiss, C. Kühnel, P. Ehrenbrink, and S. Möller. De-

velopment and validation of the conversational agents scale (cas). In *INTERSPEECH*, pages 1106–1110, 2013.

[60] S. Weech, S. Kenny, and M. Barnett-Cowan. Presence and cybersickness in virtual reality are negatively related: a review. *Frontiers in psychology*, 10:158, 2019.

[61] B. Weiss, I. Wechsung, C. Kühnel, and S. Möller. Evaluating embodied conversational agents in multimodal interfaces. *Computational Cognitive Science*, 1(1):1–21, 2015.

[62] P. Wik and A. Hjalmarsson. Embodied conversational agents in computer assisted language learning. *Speech communication*, 51(10):1024–1037, 2009.

[63] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.

[64] X. Xia, C.-M. Pun, D. Zhang, Y. Yang, H. Lu, H. Gao, and F. Xu. A 6-dof telexistence drone controlled by a head mounted display. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1241–1242. IEEE, 2019.