How to Observe Users' Movements in Virtual Environments: Viewpoint Control in a Power Wheelchair Simulator

Abdulaziz Alshaer, Umm Al-Qura University, Makkah, Saudi Arabia,
David O'Hare, Otago University, Dunedin, New Zealand,
Philippe Archambault, McGill University, Montreal, Quebec, Canada,
Mark Shirley, Southern Rehab, Dunedin, New Zealand, and
Holger Regenbrecht, Otago University, Dunedin, New Zealand

Objective: We describe a networked, two-user virtual reality (VR) power wheelchair (PWC) simulator system in which an actor (client) and an observer (clinician) meet. We then present a study with 15 observers (expert clinicians) evaluating the effect of three principal forms of viewpoint control (egocentric-egomotion, egocentric-tethered, and client-centric) on the observer's assessment of driving tasks in a virtual environment (VE).

Background: VR allows for the simulation and assessment of real-world tasks in a controlled, safe, and repeatable environment. Observing users' movement behavior in such a VE requires appropriate viewpoint control for the observer. The VR viewpoint user interface should allow an observer to make judgments equivalent or even superior to real-world situations.

Method: A purpose-built VR PWC simulator was developed. In a series of PWC driving tasks, we measured the perceived ease of use and sense of presence of the observers and compared the virtual assessment with real-world "gold standard" scores, including confidence levels in judgments.

Results: Findings suggest that with more immersive techniques, such as egomotion and tethered egocentric viewpoints, judgments are both more accurate and more confident. The ability to walk and/or orbit around the view significantly affected the observers' sense of presence.

Conclusion: Incorporating the observer into the VE, through egomotion, is an effective method for assessing users' behavior in VR with implications for the transferability of virtual experiences to the real world.

Application: Our application domain serves as a representative example for tasks where the movement of users through a VE needs to be evaluated.

Keywords: user observation, virtual reality, interaction techniques, immersion, driving simulator

Address correspondence to Holger Regenbrecht, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand; e-mail: holger.regenbrecht@otago.ac.nz.

HUMAN FACTORS
Vol. XX, No. X, Month XXXX, pp. 1–15
DOI: 10.1177/0018720819853682

INTRODUCTION

Virtual reality (VR) techniques have been increasingly used to simulate real-world behavior. In particular, for the training and assessment of potentially risky or hard-to-learn skills and behavior, VR can be a very effective and efficient instrument (Johnson, Guediri, Kilkenny, & Clough, 2011). When it comes to assessing the learnt skills, an assessor observes the user (actor) in action to judge their skills. An accurate judgment in the virtual environment (VE) is required to make sure that the observed behavior will correspond appropriately with the real world. What then is the best way to support an observer in a co-operative VR actor–observer system? Typical actor–observer systems range from a variety of vehicle simulators to observing users' movement and locomotion behavior (e.g., physiotherapy, marketing, or geo-information).

A subclass of vehicle simulators are power wheelchair (PWC) simulators designed for more effective and efficient training and assessment of driving skills. When it comes to the assessment of the newly developed skills, an assessor observes the user in action to judge the driving skills using standardized scales (Shechtman, Classen, Awadzi, & Mann, 2009). Precise judgment in the VE is essential to ensure that the observed driving behavior translates to accurate driving in the real world (Stanney, Mollaghasemi, Reeves, Breaux, & Graeber, 2003; Wang, 2001). Transferability thus becomes a key aspect of the VE. The ability to independently drive a vehicle, and, in particular, to effectively operate a PWC, makes a significant difference to the user's ability to operate in their surrounding environment. As with any other vehicle, those skills have to be learnt and assessed (Lee, 2014), and VR can potentially make a significant difference here.

A typical issue with actor–observer applications in VE settings is that observers have to perceive the simulation from the viewpoint of the actor, as a "passive observer" (Burigat & Chittaro, 2016). The observer experiences the simulator in a different way from a real-world situation, where they could freely adopt different viewpoints around the user. Observers find it difficult to mentally match the simulator situation with the real-world experience (Larsson, Västfjäll, & Kleiner, 2001). This difficulty with transferability leads to a number of problems: (1) comprehension problems, (2) poor decisions, (3) inaccurate judgments, and (4) insufficient feedback to users. Incorrect viewpoint control could lead to misperceptions of the simulation space, resulting in erroneous judgments (Alshaer, Regenbrecht, & O'Hare, 2017; Henry & Furness, 1993; Sun, Li, Zhu, & Hsiao, 2015).

Slater, Howell, Steed, Pertaub, and Garau (2000) examined how VR could be used by actors and directors to rehearse theatrical performances. The results showed that the users' sense of presence and degree of cooperation increased over time and that, ultimately, the virtual rehearsal successfully transferred to the real-world performance in a way that could not have been achieved by simply memorizing lines. There is some evidence that a first-person perspective combined with egomotion (i.e., the self-controlled movement through an environment) improves the accuracy of an observer's actions (Ogaki, Kitani, Sugano, & Sato, 2012), as well as the sense of presence (Wickens, Hollands, Banbury, & Parasuraman, 2015).

Sense of presence and immersion are two important factors in the effectiveness of any VE. Sense of presence can be defined as "the subjective experience of being in one place or environment [e.g., VE], even when one is physically situated in another" (Witmer & Singer, 1998, p. 225); others also described it as "a state of consciousness, the (psychological) sense of being in the virtual environment" (Slater & Wilbur, 1997, p. 606). When users rate high levels of presence in the VE, this results in perceiving this environment as a more engaging reality than the surrounding real world, and users consider the environment as a place they have visited rather than just images they have seen (Slater & Wilbur, 1997).

The sense of presence can be affected by hardware issues (immersion) subsequently decreasing the effectiveness of the VE (Pallavicini et al., 2013). Examples of immersion factors can be the display types, field of view, self-avatar, and/or viewpoints (Alshaer et al., 2017). Immersion can be defined as "the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant" (Slater & Wilbur, 1997, p. 605). It can be evaluated on a spectrum, from non-immersive to fully immersive (Ogle, 2002). To distinguish between presence and immersion, Schubert, Friedmann, and Regenbrecht (2001) explained that presence involves the user's experience of being part of the VE, whereas immersion involves the fidelity of the technologies used in the VE. Typically, having greater levels of immersion would result in higher levels of a sense of presence, and thus likely to create stronger psychological reactions (North & North, 2016).

Different viewpoint techniques have been developed and studied in the past. The most widely used techniques in VEs are egocentric and exocentric, and, more recently, tethered, which integrates information from both egocentric and exocentric viewpoints (Colquhoun, 2000; Jung et al., 2014). An egocentric viewpoint involves the user seeing the VE from their own first-person perspective. An exocentric view describes a viewpoint outside of first-person viewing. With a tethered viewpoint, the virtual camera (observer's viewpoint) is "attached" to the observed object or subject. Each technique has shown different effects; for example, tethered or egocentric viewpoints resulted in better performance for travel tasks (McCormick, Wickens, Banks, & Yeh, 1998); a tethered viewpoint better suits tasks that involve understanding the relations of close objects in the VE to one's own location (Hollands & Lamb, 2011). Egocentric viewpoints provide greater sense of self when compared with exocentric viewpoints (Ma & Kaber, 2006), and users acquired spatial knowledge more effectively when using active navigation (egomotion) compared with passive navigation (Burigat & Chittaro, 2016).

What has not yet been investigated is the question of which viewpoint control is most appropriate for an observer in contrast to the actor navigating through (virtual) space. In our case, what is the best viewpoint control user interface for an observer to judge the performance of a PWC user? The evaluation task requires the observer to follow the PWC, to observe the user from different angles (Hafid & Inoue, 2005). This assessment activity requires the active observation of involved observers (Burigat & Chittaro, 2016; Hafid & Inoue, 2005; Hughes & Lewis, 2005).

In this paper, we present a networked, two-user VR PWC simulator, in which a user (actor) and an assessor (observer) meet for evaluation purposes. We compared three different forms of viewpoint control: (1) egocentric-egomotion, where the observer walks around the driving user; (2) egocentric-tethered, where the observer orbits through the virtual scene around the driving user (using a standard mouse); and (3) a client-centric viewpoint, which is commonly used in PWC simulators. Fifteen observers (expert clinicians) evaluated the impact of these viewpoints on the assessment of driving tasks in a VE. Henceforth, the viewpoints will be identified as "walk" (egocentric-egomotion using a head-mounted display [HMD]), "orbit" (egocentric-tethered using a standard mouse), and "standard" (client-centric using a desktop PC interface).

Our research addresses the following questions:

Research Question 1: Do different viewpoints affect how an observer assesses driving in the VE, resulting in different judgments?

Research Question 2: How does each viewpoint affect the observer's sense of presence and confidence level?

Research Question 3: How can observers (e.g., clinicians) validly assess driving tasks in the VE compared with pre-assessed real-world driving tasks?

We hypothesize that more immersive viewpoints will lead to more valid judgments, a higher confidence level, and a greater sense of presence.

METHOD

The Wheelchair Skills Test (WST) was used as a reference for designing and developing the driving tasks (Kirby, Swuste, Dupuis, MacLeod, & Monroe, 2002). The WST evaluates PWC driving skills by means of a comprehensive scoring system where users receive a score: 0 (*fail*), 1 (*pass with difficulty*), or 2 (*pass*) for each task/skill. Only few studies on PWC simulators have incorporated clinical assessment (Kamaraj, Dicianno, Mahajan, Buhari, & Cooper, 2016; Mahajan, Dicianno, Cooper, & Ding, 2013). These studies suffer from common experimental design issues: (1) the same observer assessed the same user doing the same tasks in all conditions, which could lead to similar judgments (scores) for all conditions; (2) the participant's driving skills would inevitably improve by repeating the same tasks over and over again (even with randomization); (3) the driving assessments were based on actor-centric information, thus suffering from a "locked" frame of reference that clinicians do not face in real-world scenarios.

Our study was designed to avoid, or at least mitigate, the aforementioned design problems. First, four tasks with varying difficulties were selected from the WST out of 32 (Figure 1). The WST scoring system is used to model the selected driving tasks in the real world by an expert clinician (from the RATA Southern Rehab Clinic, Dunedin, New Zealand) and an expert PWC user. Second, the clinician gave the PWC user (a volunteer PWC user with mobility issue) instructions on how to complete each task, as a 0 (*fail*), 1 (*pass with difficulty*), or 2 (*pass*). This resulted in a total of 12 modeled driving tasks.

Third, the driving tasks of the expert PWC user were recorded by tracking the movements of the PWC user prior to the experiment. In real time, movements were animated and then saved in the simulator (Figure 2).

Fourth, three independent professional therapists (from the Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Laval University, Quebec City, Canada) evaluated the recorded tasks to assure the

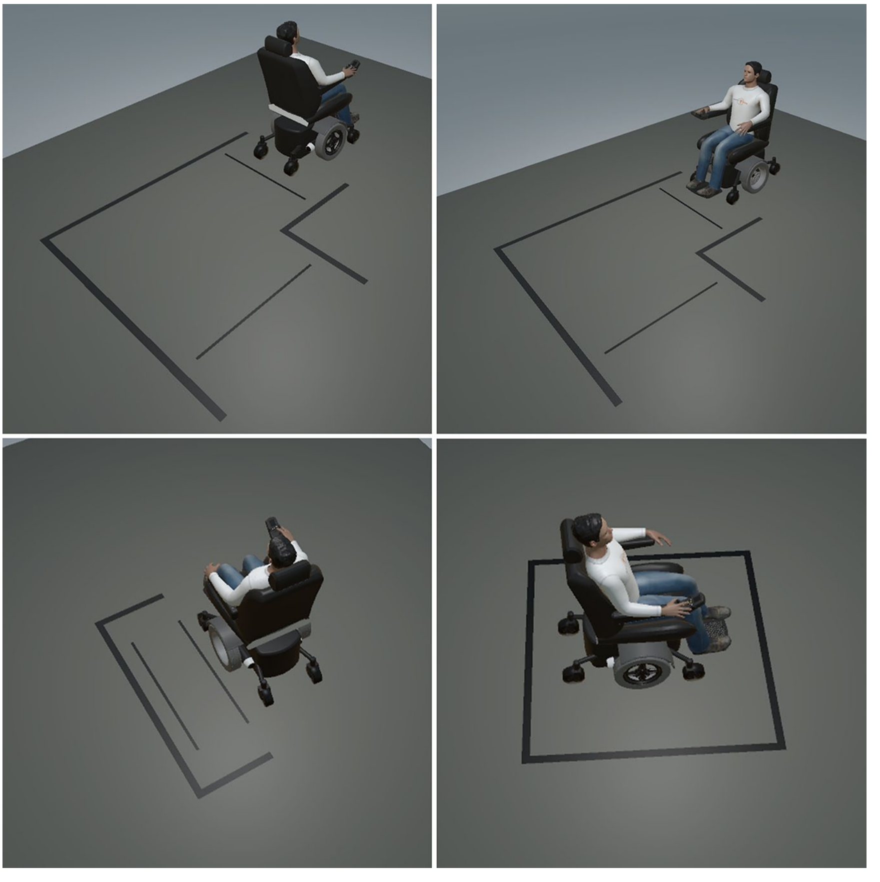

Figure 1. Illustration of the selected driving task based on WST. Driving backward (top left), driving forward (top right), sideways maneuver (bottom left), and turn 180° (bottom right). WST = Wheelchair Skills Test.

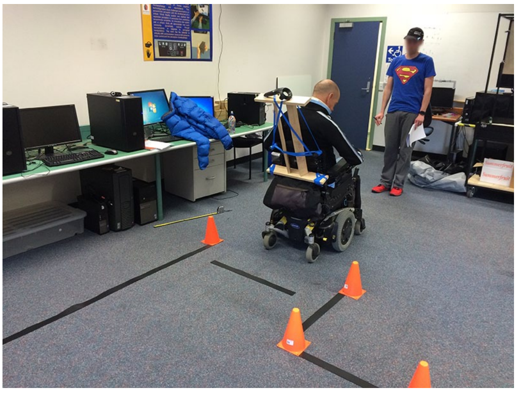

Figure 2. Virtual reality HTC controller assembled on a wooden frame to track the PWC movement in real time. The X and Y positions and orientations of the controller were tracked within the environment. PWC = power wheelchair.

validity of the assessment score originally assigned by the expert clinician. Fifth, the recorded tasks were then re-played for clinicians to assess their capacity levels as 0 (*fail*), 1 (*pass with difficulty*), or 2 (*pass*) which were then compared with the original score. Figure 3 shows an overview of the study design.

Participants

Participants were recruited from the Jewish Rehabilitation Hospital (CISSS Laval) in Montreal, Canada. Fifteen expert clinicians—eight physiotherapists (PT) and seven occupational therapists (OT)—took part in the experiment. There were four male and 11 female clinicians, with an average age of 34.9 years (*SD* = 9.4, age range = 23–55) and an average working experience of 9.3 years (*SD* = 7.73). Eight

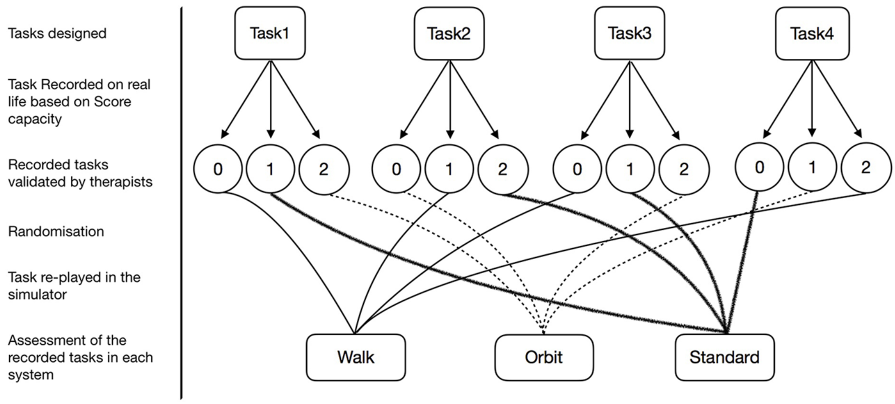

Figure 3. Overview of the study design.

Figure 4. Walk condition (top) and orbit condition (bottom).

clinicians had experience with PWC assessment with an average of 1.6 years. All participants were rewarded with chocolate bars. Institutional ethical approvals were obtained from McGill University. Figure 4 shows participants performing the task.

Apparatus

The display components of the system (Figure 5) consisted of two screens (24 in.) and a VR HMD (HTC Vive). Participants in the orbit condition used a standard mouse to orbit around the virtual PWC. In the walk condition, the HTC Vive headset was used by the participants to walk around the virtual PWC (Figure 6).

Environment

The VE in this experiment consists of the virtual PWC, an actor's avatar, and the driving tasks themselves. The virtual PWC was modeled on the real PWC that was used to record the driving tasks. In the orbit condition, clinicians were allowed to set their preferred distance from the virtual PWC using the mouse's scroll wheel while being able to change the point of view by moving the mouse around. The sphere in which the orbiting viewpoint moved around was limited from both the top and bottom (as can be seen in Figure 7). This was done to ignore any extra drag in the mouse and to avoid spinning around the virtual PWC. To ensure the user's safety, precautions were taken to stop the user from going beyond the limited space in the walk condition, by showing red boundaries in the virtual environment once the user got within 50 cm of the boundaries.

Experimental Design

The design was a 3 (viewpoints) × 4 (virtual driving tasks) within-subjects factorial design. This yielded 12 conditions as shown in Table 1.

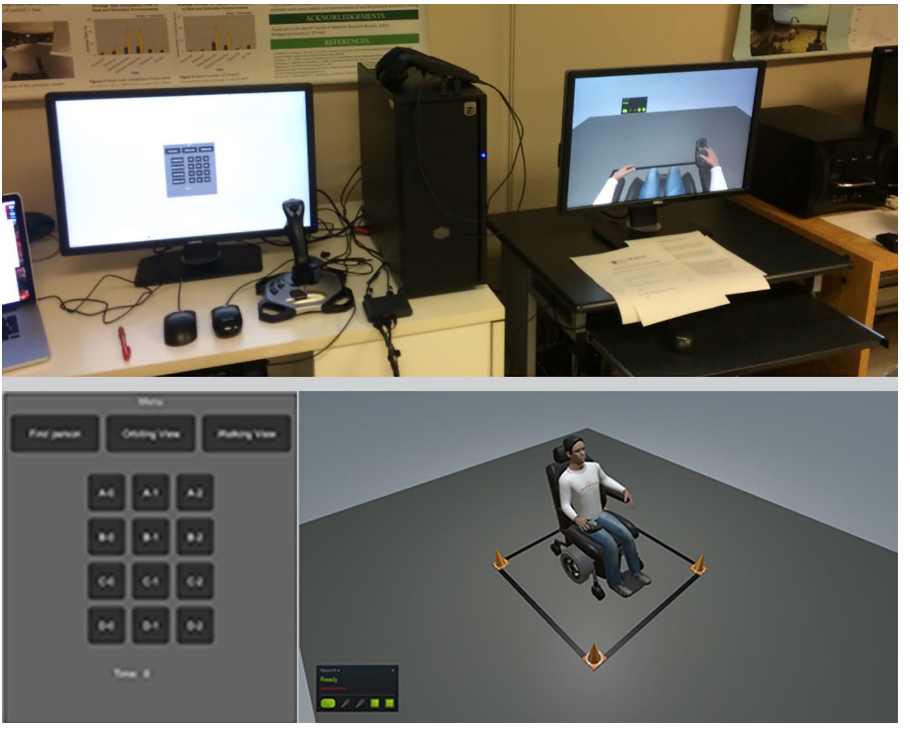

Figure 5. (Top) the two screens. (Bottom left) the experimenter view to change the viewpoints and the driving tasks and (bottom right) a screenshot of the orbiting view.

Figure 6. For the walk condition, a space of 4 × 4 m^2 was cleared and tracked. The stationary trackers ("lighthouses") were mounted on two separate tripods at 2.5 m height.

Measured variables include assessment score, ease, confidence level, and sense of presence. The driving tasks included the following: turn while moving forward (T_1), turn while moving backward (T_2), turn in place (T_3), and sideways maneuvers (T_4).

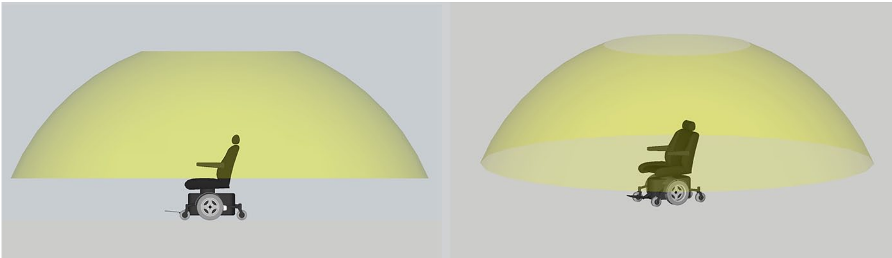

Figure 7. The yellow sphere shows where clinicians can orbit around the virtual PWC with the ability to proximity at any given time. PWC = power wheelchair.

Measures

Assessment score. Assessment scores were based on the WST. The scoring system consisted of three levels: 0 ("task incomplete or unsafe"), 1 ("evaluation criteria are met but the subject experienced some difficulty worthy of note"), and 2 ("task independently and safely accomplished without any difficulty"). To analyze the virtual assessment score and compare it with the original score from the real-world designed driving tasks (correct score), the square of the difference score (Correct Score – Judged Score) was calculated. Scores were entered into a factorial analysis of variance (ANOVA) with condition (standard, orbit, walk) and task (T1, T2, T3, T4) as within-subjects factors. The number of correct answers (matching the assigned score) was also analyzed.

The confidence level expressed how sure the clinicians were about the accuracy of their assessment score. Clinicians were asked to rate their confidence level on a 7-point Likert-type scale question after assessing each task ("Confidence level when the task was assessed: 1 = *very uncertain*, 7 = *very certain*"). For sense of presence, the Igroup Presence Questionnaire (IPQ) was used (Schubert et al., 2001). The IPQ questionnaire consists of 13 questions and defines the user's general sense of presence, involvement, spatial presence, and realism; each question took the form of a 7-point scale after each condition, for example, "In the computer generated world, I had a sense of 'being there.'"

Counterbalancing. To mitigate potential learning effects that could arise from assessing the same tasks in all three conditions, the following measures were taken: (1) the condition order was randomized in a counterbalanced order, and (2) within each subject, the score of the tasks was different in each condition (e.g., T_1 would be "0" in the first condition, "1" in the second condition, and "2" in the third condition regardless of the condition order). In this case, the subject assessed four different tasks in each condition. In addition, the order of the tasks represented in each condition was randomized, which made it difficult for participants to guess the task score in the third condition. Table 2 shows the randomization of the conditions, tasks, and task level.

TABLE 1: A 3 × 4 Factorial Design

	Driving Tasks			
Viewpoints	T_1	T_2	T_3	T_4
Standard	S-T_1	S-T_2	S-T_3	S-T_4
Orbit	O-T_1	O-T_2	O-T_3	O-T_4
Walk	W-T_1	W-T_2	W-T_3	W-T_4

Participant's Task

The participant's (clinician's) task was to watch pre-recorded PWC driving tasks and assess them based on the WST. In the standard condition, the participant's task was to sit down, watch the driving task (from the perspective of the PWC user), and assign a score at the end. In the orbit condition, the participant's task was to use the mouse to change the proximity to the wheelchair, orbit around the driving tasks, and assign a score at the end. In the walk condition, the participant's task was to use the HMD, walk around the recorded driving task, and assign a score at the end.

TABLE 2: A Block of Complete Randomization Repeated by Every Three Subjects (T_1-0 Means Task 1—Correct Scores of 0)

Subject 1	Standard				Orbit				Walk			
	T_1-0	T_4-2	T_2-1	T_3-0	T_3-1	T_2-2	T_4-0	T_1-1	T_4-1	T_3-2	T_1-2	T_2-0
Subject 2	Walk				Standard				Orbit			
	T_1-0	T_4-2	T_2-1	T_3-0	T_3-1	T_2-2	T_4-0	T_1-1	T_4-1	T_3-2	T_1-2	T_2-0
Subject 3	Orbit				Walk				Standard			
	T_1-0	T_4-2	T_2-1	T_3-0	T_3-1	T_2-2	T_4-0	T_1-1	T_4-1	T_3-2	T_1-2	T_2-0

TABLE 3: Virtual Assessment Means and Standard Deviations

	Driving Tasks				
Viewpoints	T_1	T_2	T_3	T_4	Average
Standard	0.6 (0.13)	1.1 (0.41)	0.7 (0.31)	0.5 (0.27)	0.73
Orbit	0.13 (0.1)	0.33 (0.13)	0.2 (0.11)	0.27 (0.12)	0.23
Walk	0.07 (0.07)	0.13 (0.1)	0.00 (0.00)	0.2 (0.11)	0.1
	0.27	0.51	0.31	0.33	

Procedure

Upon arrival, observers were welcomed, given the information sheet to read, and then asked to sign the consent form and fill out a demographics questionnaire. Participants were given the opportunity to try each setup for as long as they needed with pre-recorded driving tasks for this purpose. The experimental procedure was then explained to the participants, including the condition order and the nature of the driving tasks. Participants were given the driving task description sheet and the assessment score criteria.

Then, participants were asked to read the assessment questionnaire and told they would verbally answer the questions after each task. These questions included the following: (1) "Based on the Wheelchair Skills Test (WST) Version 4.2, I give this driving task a capacity score of X" (3-point scale); (2) "Using this setup, the assessment of this task was X" (7-point scale); and (3) "Confidence level when the task was assessed" (7-point scale). Participants were given the sense-of-presence questionnaire. Finally, participants were debriefed and given a chocolate bar. The entire procedure took approximately 40 min per participant.

RESULTS

Assessment Scores

Virtual assessment scores were compared with the correct scores based on the pre-recorded driving tasks. The means and standard deviations of the assessment scores are reported in Table 3. The walk condition showed the lowest difference between correct and judged scores (M = 0.1, SD = 0.04) followed by orbit (M = 0.23, SD = 0.5) and standard condition (M = 0.73, SD = 0.14). ANOVA showed no significant interaction between viewpoints and tasks on the clinicians' assessment. There was also no significant mean effect for the tasks. However, ANOVA confirmed a significant viewpoints main effect on clinician assessment, $F(2, 28) = 14.1$, $p < .001$, $\omega^2 = 0.5$. A post hoc test (Tukey's) showed that the standard condition differed significantly from both orbit (p = .013) and walk conditions (p = .001). There was no significant difference between orbit and walk conditions.

Perceived Ease of Use

Means and standard deviations are reported in Table 4. It can be seen that the orbit and walk conditions were rated the easiest to assess

TABLE 4: Means and Standard Deviations for Ease of Use Question

	Driving Tasks				
Viewpoints	T1	T2	T3	T4	Average
Standard	4.5 (0.3)	3 (0.43)	4.5 (0.43)	3.9 (0.44)	3.96
Orbit	5.5 (0.35)	6.2 (0.17)	6 (0.36)	6.3 (0.25)	6.01
Walk	5.7 (0.42)	5.9 (0.32)	6.1 (0.32)	6.4 (0.34)	6.02
	5.22	5.02	5.55	5.53	

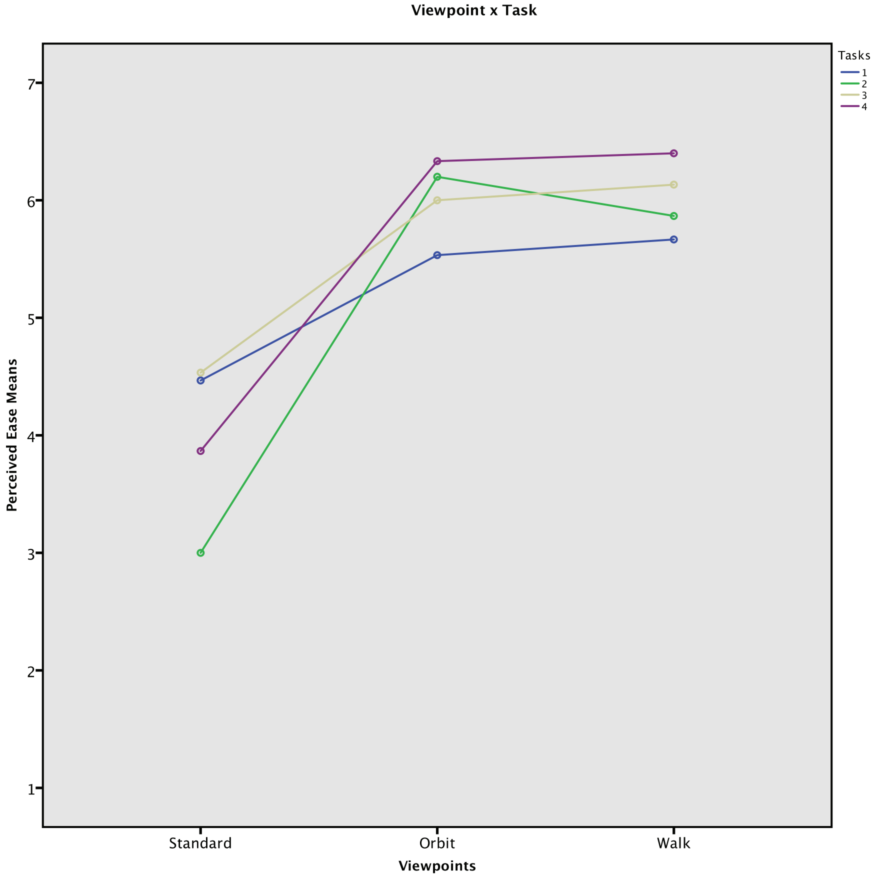

Figure 8. Interaction between viewpoints and tasks for perceived ease.

each task compared with the standard condition. ANOVA confirmed significant interaction effects between viewpoint and task, $F(3.06, 84) = 3.3$, $p < .006$, $\omega^2 = 0.19$ (Figure 8). Significant main effects were also revealed for both viewpoints, $F(2, 28) = 19.63$, $p < .001$, $\omega^2 = 0.58$, and tasks, $F(3, 42) = 3.76$, $p < .018$, $\omega^2 = 0.21$. A post hoc test (Tukey's) showed that the standard viewpoint differed significantly from both orbit and walk viewpoints ($p = .001$). It also showed that T_2 (moving backward task) was perceived to be significantly more difficult to assess than T_3 (turn in place task) and T_4 (sideways maneuver task) with $p = .046$ and $p = .016$ (respectively).

TABLE 5: Means and Standard Deviations for Confidence Level

	Driving Tasks				
Viewpoints	T_1	T_2	T_3	T_4	Average
Standard	4.8 (0.3)	3.2 (0.43)	4.4 (0.4)	4.2 (0.46)	4.15
Orbit	5.3 (0.34)	6.06 (0.2)	6.06 (0.35)	6 (0.29)	5.86
Walk	6.1 (0.28)	5.6 (0.29)	6.1 (0.4)	6.4 (0.25)	6.03
	5.4	5.0	5.5	5.53	

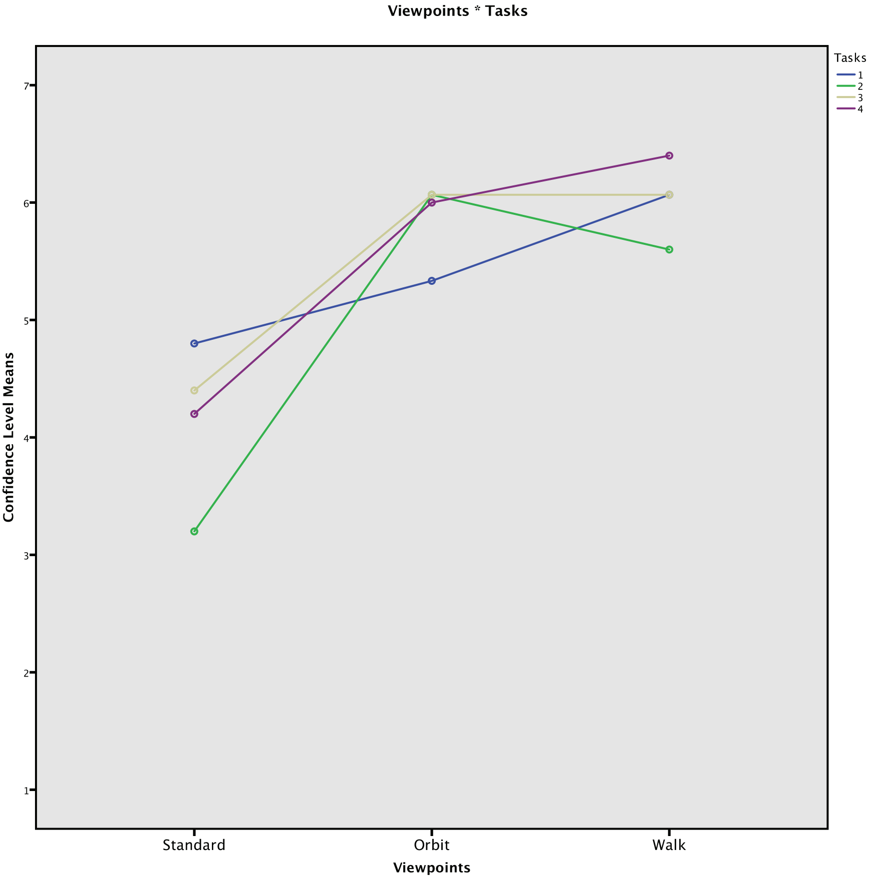

Figure 9. Interactions between viewpoints and tasks for confidence level.

Confidence Level

Means and standard deviations are reported in Table 5. An ANOVA of the confidence level revealed a significant interaction between viewpoints and tasks, $F(6, 84) = 2.94$, $p < .012$, $\omega^2 = 0.174$ (see Figure 9). ANOVA also revealed a significant viewpoints main effect on the clinicians' confidence level, $F(2, 28) = 21.3$, $p < .001$, $\omega^2 = 0.6$, with the walk condition being the highest ($M = 6.03$, $SD = 0.23$), followed by orbit ($M = 5.87$, $SD = 0.24$) and standard conditions ($M = 4.15$, $SD = 0.26$). Although means differ slightly between tasks, there were no significant differences.

Correct Answers

Correct answers were based on the number of judged scores accurately matched to correct

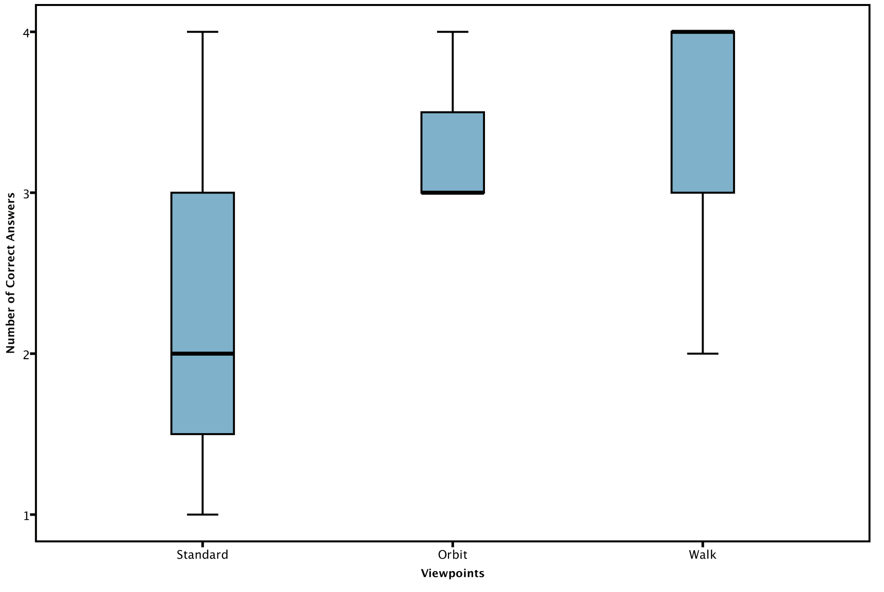

Figure 10. Number of correct answers. Box plot represents the median, interquartile (blue box), minimum, and maximum.

TABLE 6: Means and Standard Deviations for General Sense of Presence, Involvement, Spatial Presence, and Realism (Scale –3 to 3)

	Standard	Orbit	Walk
General sense of presence	–1.8 (1.14)	–0.4 (1.68)	2.6 (0.8)
Involvement	–0.3 (1.04)	–0.3 (0.87)	1.1 (0.9)
Spatial presence	–2.3 (0.72)	–0.96 (1.34)	2.4 (0.93)
Realism	–1.8 (1.12)	–0.77 (1.15)	0.3 (0.98)

scores. A one-way ANOVA was conducted to compare the effect of the viewpoints on the participant's judgment. The overall number of correct answers (out of 60) rose from 34 (standard) to 42 (orbit) and 54 (walk). Participants made more correct answers when using the walk condition (M = 3.53, SD = 0.64), followed by orbit (M = 3.07, SD = 0.7) and standard (M = 2.27, SD = 1.03). Figure 10 shows boxplots representing the correct answer responses. This was confirmed by the results of the statistical analysis (one-way ANOVA), which showed that the viewpoints had a significant main effect on the number of correct answers, $F(2, 42) = 9.36$, $p = .001$. A post hoc test showed that there was a significant main effect between standard and orbit ($p = .030$) and standard and walk ($p = .001$) conditions. However, there was no significant effect between the orbit and walk conditions.

Sense of Presence

For sense of presence, feedback was collected on general sense of presence, involvement, spatial presence, and realism. The means and standard deviations are reported in Table 6. Overall, the walk condition was rated the highest in terms of overall general sense of presence, involvement, spatial presence, and realism; all means were above the mid-point, whereas the means of the standard and orbit conditions were below the mid-point (see Figure 11).

For the general sense of presence, the effects of the viewpoints on the participants' sense of presence were significant, $F(2, 42) = 47.09$,

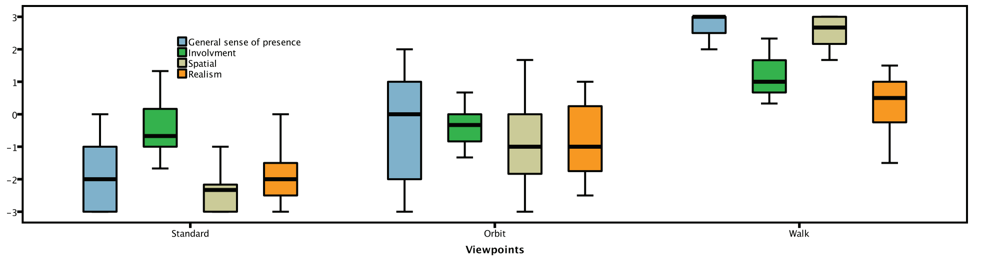

Figure 11. Boxplots representing all sense of presence categories. Box plots represent the median, interquartile (colored boxes), minimum, and maximum.

$p < .001$. A post hoc test (Tukey's) showed significant differences between all the conditions: between the standard and orbit, $p < .013$; between the standard and walk conditions, $p < .001$; and between the orbit and walk conditions, $p = .001$. For involvement, the effects of the viewpoints on user involvement were also significant, $F(2, 42) = 11.1$, $p < .001$. A post hoc test showed significant differences between standard and walk conditions, $p < .001$, and between orbit and walk conditions, $p < .001$. There was no significant difference between the standard and orbit conditions.

For spatial presence, the effects of the viewpoint on user involvement were significant, $F(2, 42) = 81.1$, $p < .001$. A post hoc test showed significant differences between all the conditions: between the standard and orbit conditions, $p = .003$; between the standard and walk conditions, $p < .001$; and between the orbit and walk conditions, $p = .001$. For realism, the effects of the viewpoint on user realism were significant, $F(2, 42) = 14.4$, $p < .001$. A post hoc test showed significant differences between all the conditions: between the standard and orbit conditions, $p = .031$; between the standard and walk conditions, $p < .001$; and between the orbit and walk conditions, $p = .031$.

DISCUSSION AND CONCLUSION

This study investigated three forms of viewpoint control (egocentric-egomotion, egocentric-tethered, and client-centric viewpoints) for an observing assessor in a PWC driving simulator. There was a strong main effect of egocentric viewpoints on participants' assessment, perceived ease of use, and confidence level compared with client-centric views. For all three analyses, the effect sizes were large to very large. Although egomotion had a mixed impact (for instance, it did not significantly improve judgmental accuracy compared with the orbit but did increase sense of presence), the frame of reference makes a clear difference to the participants' judgment. Although there was significant interaction between tasks and viewpoints for participants' perceived ease of use and confidence level, neither tasks nor interaction between tasks and viewpoints had any main effect on observers' judgment scores.

Assessment and Correct Answers

The walk condition (egocentric viewpoint) was the most effective form for virtual assessment with respect to real-world scores. This supports the findings by Burigat and McCormick (Burigat & Chittaro, 2016; McCormick et al., 1998) where users acquired spatial knowledge more effectively when using active navigation. Unlike the orbit and standard view, the difference between the judged and correct scores when using the walk condition was minimal and, in some tasks, was even zero. The selected tasks varied in their difficulties from both the user's perspective (to drive) and the observer's perspective (to assess); yet, they had no effect on the observers' judged scores.

The use of the orbit view in the experiment was to provide observers with an alternative low-cost solution to the walk condition. With the orbit technique, only a standard desktop, monitor, and mouse are required, whereas the walk technique requires a more expensive HMD, a large space, a complex setup, and a powerful desktop computer. Observers were also able to make significantly better judgments when using the orbit view, compared with the standard condition. This could be due to the fact that tethered viewpoint helped observers to better understand the relations of the virtual PWC to their own location as suggested by Hollands and Lamb (2011).

The results of this experiment offer evidence for the importance of the viewpoint perspective. For example, although the walk and orbit conditions vary in terms of egomotion, both are observer-centered points of view. Post hoc comparisons showed no differences between these conditions on judgmental accuracy. On the contrary, although the orbit and standard conditions vary in terms of point of view, neither involve egomotion. Post hoc comparisons showed significant differences on the score accuracy and the number of correct answers between these conditions. These findings suggest that the frame of reference makes a significant difference to judgment, but that physically moving through the virtual space does not. This finding contradicts the findings of Wickens et al. (2015) who showed better accuracy when first-person perspective was combined with egomotion (self-movements). However, the addition of egomotion does make a difference to the experienced sense of presence similar to Ma and Kaber (2006) and Wickens et al. (2015). The standard condition may have provided the clinicians with different insights into the driving tasks, as they viewed it from the perspective of a PWC user, but it visually restricted the clinician's viewpoint and made the assessment more difficult.

Perceived Ease and Confidence Level

The viewpoints significantly affected observers in terms of both how easy each system was to evaluate the tasks, and their confidence level when assessing the driving tasks. The results showed that the more the system was perceived as easy to use, the higher the confidence level. For example, there was a 50% increase in confidence level for both the orbit and walk conditions. The significant increase of the confidence level over the standard condition shows the advantage of incorporating viewpoint interaction within the simulator.

Sense of Presence

Overall, the walk condition was rated the highest across all conditions. Only the involvement aspect was not significantly different between the standard and orbit conditions. Although the walk condition was rated significantly higher than the orbit condition in all the sense-of-presence factors (general sense of presence, realism, spatial presence, and involvement), this did not impact perceived ease, confidence level, or the number of correct answers. These findings suggest that even with a less immersive simulator, clinicians could still make accurate judgments.

Application

This paper presented a system and experiment to evaluate the effect of three forms of viewpoint control (walk, orbit, and standard) for actor–observer systems. The results showed that the walk and orbit condition allowed observers to make more accurate and more confident judgments.

One of the limitations of this study was the restricted tracked space where the observer could move which might lead to the perception of egocentric distances to be more compressed (Interrante, Ries, & Anderson, 2006; Rousset, Bourdin, Goulon, Monnoyer, & Vercher, 2015; Sinai, Krebs, Darken, Rowland, & McCarley, 1999; Willemsen & Gooch, 2002). We assume the effect will be minimal for this study, as the observer has standardized scales to base judgments on.

The findings and insights gained from this study can be applied to a range of observer–actor simulators, where a supervisor, observer, or assessor is involved in training and/or evaluation processes. Extending the navigation techniques presented and evaluated in this study would allow for applications and investigations for mobility scooters, vessels, and other vehicles, as well as applications where users walk around VEs for observation purposes. For all future applications and experiments, an embodied viewpoint control for the observer would lead to more accurate judgments of user performance.

ACKNOWLEDGMENTS

The authors wish to thank all of the therapists who participated in this experiment, and the HCI group at the University of Otago. Thanks must also go to François Routhier, Josianne Lettre, and Krista Best for their help with the assessment of the recorded driving tasks. A special thanks goes to Catherine Bigras for taking part in the experiment as a translator, and Keith Jarvie for helping with the recording of the driving tasks. The first author is sponsored by the Saudi Arabian Ministry of Education and Umm Al-Qura University (NZ090182). Abdulaziz Alshaer is also affiliated with Umm Al-Qura University, Makkah, Saudi Arabia.

KEY POINTS

- We presented a purpose-built, networked, virtual reality power wheelchair simulator as a representative example for systems where users' movements through virtual environments need to be evaluated.
- Observers' ability to judge the driving behavior of users is more effective with more immersive techniques of viewpoint control, such as actual ego-motion and tethered orbiting.
- Our findings have implications for the transferability of virtual experiences to the real world for wheelchair and other simulators.

ORCID IDS

David O'Hare https://orcid.org/0000-0002-5156-3925

Holger Regenbrecht https://orcid.org/0000-0002-5037-6407

REFERENCES

Alshaer, A., Regenbrecht, H., & O'Hare, D. (2017). Immersion factors affecting perception and behaviour in a virtual reality power wheelchair simulator. *Applied Ergonomics*, *58*, 1–12.

Burigat, S., & Chittaro, L. (2016). Passive and active navigation of virtual environments vs. traditional printed evacuation maps: A comparative evaluation in the aviation domain. *International Journal of Human-Computer Studies*, *87*, 92–105.

Colquhoun, H. W. (2000). *Dynamic tethering for enhanced remote control and navigation*. Ottawa, Ontario: National Library of Canada.

Hafid, N., & Inoue, T. (2005). *Electric wheelchair simulator for rehabilitation of persons with motor disability*. Paper presented at the Simpósio Brasileiro De Realidade Virtual, Belem, Brazil.

Henry, D., & Furness, T. (1993). Spatial perception in virtual environments: Evaluating an architectural application. In *Proceedings of the IEEE Virtual Reality Annual International Symposium* (pp. 33–40). New York, NY: Institute of Electrical and Electronics Engineers.

Hollands, J. G., & Lamb, M. (2011). Viewpoint tethering for remotely operated vehicles: Effects on complex terrain navigation and spatial awareness. *Human Factors*, *53*, 154–167.

Hughes, S., & Lewis, M. (2005). Attentive navigation for viewpoint control in virtual environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*, 630–643.

Interrante, V., Ries, B., & Anderson, L. (2006). Distance perception in immersive virtual environments, revisited. In *Proceedings of the IEEE Conference on Virtual Reality* (pp. 3–10). New York, NY: Institute of Electrical and Electronics Engineers.

Johnson, S. J., Guediri, S. M., Kilkenny, C., & Clough, P. J. (2011). Development and validation of a virtual reality simulator: Human factors input to interventional radiology training. *Human Factors*, *53*, 612–625.

Jung, J., Park, H., Hwang, D., Son, M., Beck, D., Park, J., & Park, W. (2014, January 7–9). *A review on interaction techniques in virtual environments*. Paper presented at the 2014 International Conference on Industrial Engineering and Operations Management, Bali, Indonesia.

Kamaraj, D. C., Dicianno, B. E., Mahajan, H. P., Buhari, A. M., & Cooper, R. A. (2016). Interrater reliability of the power mobility road test in the virtual reality-based simulator-2. *Archives of Physical Medicine and Rehabilitation*, *97*, 1078–1084.

Kirby, R. L., Swuste, J., Dupuis, D. J., MacLeod, D. A., & Monroe, R. (2002). The Wheelchair Skills Test: A pilot study of a new outcome measure. *Archives of Physical Medicine and Rehabilitation*, *83*, 10–18.

Larsson, P., Västfjäll, D., & Kleiner, M. (2001). The actor-observer effect in virtual reality presentations. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *4*, 239–246.

Lee, S.-H. (2014). Users' satisfaction with assistive devices in South Korea. *Journal of Physical Therapy Science*, *26*, 509–512.

Ma, R., & Kaber, D. B. (2006). Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies*, *64*, 541–552.

Mahajan, H. P., Dicianno, B. E., Cooper, R. A., & Ding, D. (2013). Assessment of wheelchair driving performance in a virtual reality-based simulator. *The Journal of Spinal Cord Medicine*, *36*, 322–332.

McCormick, E. P., Wickens, C. D., Banks, R., & Yeh, M. (1998). Frame of reference effects on scientific visualization subtasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*, 443–451.

North, M. M., & North, S. M. (2016). A comparative study of sense of presence of traditional virtual reality and immersive environments. *Australasian Journal of Information Systems*, *20*, 1–15.

Ogaki, K., Kitani, K. M., Sugano, Y., & Sato, Y. (2012). *Coupling eye-motion and ego-motion features for first-person activity recognition*. In *Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1–7). New York, NY: Institute of Electrical and Electronics Engineers.

Ogle, T. (2002). *The effects of virtual environments on recall in participants of differing levels of field dependence* (Doctoral dissertation, Virginia Polytechnic and State University, Blacksburg). Retrieved from https://vtechworks.lib.vt.edu/handle/10919/27245

Pallavicini, F., Cipresso, P., Raspelli, S., Grassi, A., Serino, S., Vigna, C., & Riva, G. (2013). Is virtual reality always an effective stressors for exposure treatments? Some insights from a controlled trial. *BMC Psychiatry*, *13*, Article 52.

Rousset, T., Bourdin, C., Goulon, C., Monnoyer, J., & Vercher, J. L. (2015). Does virtual reality affect visual perception of egocentric distance? In *Proceedings of the 2015 IEEE Virtual Reality (VR)* (pp. 277–278). New York, NY: Institute of Electrical and Electronics Engineers.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, *10*, 266–281.

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: A validation study. *Traffic Injury Prevention*, *10*, 379–385.

Sinai, M., Krebs, W., Darken, R., Rowland, J., & McCarley, J. (1999). Egocentric distance perception in a virtual environment using a perceptual matching task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *43*, 1256–1260.

Slater, M., Howell, J., Steed, A., Pertaub, D.-P., & Garau, M. (2000). Acting in virtual reality. In *Proceedings of the Third International Conference on Collaborative Virtual Environments* (pp. 103–110). New York, NY: Association for Computing Machinery.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *6*, 603–616.

Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., & Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): Identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies*, *58*, 447–481. doi:10.1016/S1071-5819(03)00015-6

Sun, H.-M., Li, S.-P., Zhu, Y.-Q., & Hsiao, B. (2015). The effect of user's perceived presence and promotion focus on usability for interacting in virtual environments. *Applied Ergonomics*, *50*, 126–132.

Wang, W. (2001). Dynamic viewpoint tethering: Controlling a virtual camera for effective navigation in virtual environments. In *Proceedings of the CHI'01 Extended Abstracts on Human Factors in Computing Systems* (pp. 93–94). New York, NY: Association for Computing Machinery.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering Psychology and Human Performance*. New York, NY: Psychology Press.

Willemsen, P., & Gooch, A. A. (2002). Perceived egocentric distances in real, image-based, and traditional virtual environments. In *Proceedings of the IEEE Virtual Reality* (pp. 275–276). New York, NY: Institute of Electrical and Electronics Engineers.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*, 225–240.

Abdulaziz Alshaer is an assistant professor in the department of Computer Science at Umm Al-Qura University. He received his PhD from Otago University in software engineering in 2017. His research interests include HCI, virtual reality, and information technology.

David O'Hare is a professor in the Department of Psychology at the University of Otago. He received his PhD in 1978. His research interests include decision making under uncertainty, case-based learning, display design, and flight simulation.

Philippe Archambault is an occupational therapist and associate professor at the School of Physical and Occupational Therapy, McGill University. He received his PhD from Montreal University in neuroscience in 2003. His research focuses on the use of technology in rehabilitation.

Mark Shirley is a physiotherapist and executive director of Southern Rehab. He received his master's in health science from Auckland University in 2011. His research interests include advanced technology developments and how they can assist inter-disciplinary rehabilitation.

Holger Regenbrecht is a professor in the Department of Information Science at Otago University. He received his PhD in computer science from Bauhaus University in 1999. His research interests include HCI, applied computer science and information technology, mixed reality, telepresence, and computer-aided therapy and rehabilitation.

Date received: July 15, 2018
Date accepted: April 19, 2019