Journal of Perceptual Imaging® 3(2): 020502-1–020502-10, 2020.

Contribution of Motion Parallax and Stereopsis to the Sense of Presence in Virtual Reality

Siavash Eftekharifar[1,2], Anne Thaler[2,3], Nikolaus F. Troje[1,2,3]

[1] *Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada;* [2] *Centre for Vision Research, York University, Toronto, Ontario, Canada;* [3] *Department of Biology, York University, Toronto, Ontario, Canada*
E-mail: S.Eftekharifar@queensu.ca

Abstract. *The sense of presence is defined as a subjective feeling of being situated in an environment and occupying a location therein. The sense of presence is a defining feature of virtual environments. In two experiments, we aimed at investigating the relative contribution of motion parallax and stereopsis to the sense of presence, using two versions of the classic pit room paradigm in virtual reality. In Experiment 1, participants were asked to cross a deep abyss between two platforms on a narrow plank. Participants completed the task under three experimental conditions: (1) when the lateral component of motion parallax was disabled, (2) when stereopsis was disabled, and (3) when both stereopsis and motion parallax were available. As a subjective measure of presence, participants completed a presence questionnaire after each condition. Additionally, electrodermal activity (EDA) was recorded as a measure of anxiety. In Experiment 1, EDA responses were significantly higher with restricted motion parallax as compared to the other two conditions. However, no difference was observed in terms of the subjective presence scores across the three conditions. To test whether these results were due to the nature of the environment, participants in Experiment 2 experienced a slightly less stressful environment, where they were asked to stand on a ledge and drop virtual balls to specified targets into the abyss. The same experimental manipulations were used as in Experiment 1. Again, the EDA responses were significantly higher when motion parallax was impaired as compared to when stereopsis was disabled. The results of the presence questionnaire revealed a reduced sense of presence with impaired motion parallax compared to the normal viewing condition. Across the two experiments, our results unexpectedly demonstrate that presence in the virtual environments is not necessarily linked to EDA responses elicited by affective situations as has been implied by earlier studies.*
[DOI: 10.2352/J.Percept.Imaging.2020.3.2.020502]

1. INTRODUCTION

The main goal of virtual reality (VR) devices is to create immersive virtual environments that elicit a sense of presence in users. The sense of presence (SOP) is the subjective feeling of being at a certain location [4, 8, 13, 15, 24, 25, 27, 30, 32]. We experience it in real, visual space but not when we look at pictures. Slater [27] talks about the "place illusion" as SOP is also experienced in VR. The sensation is illusory in the sense that the user is physically situated elsewhere (and being aware of it). Slater suggests that the degree of SOP in VR depends on the sensorimotor contingencies that the immersive system supports. Several depth cues seem to be involved in that respect; however, their relative contribution to eliciting a subjective feeling of presence in VR is not clear. On the one hand, there are pictorial depth cues (occlusion, linear perspective, shape from shading, texture gradient, etc.) that can be represented in static and moving pictures. However, they seem to contribute little to creating an SOP [13]. On the other hand, stereopsis and motion parallax are considered essential parts of immersive virtual environments, which critically distinguishes them from pictorial renderings. While binocular disparities are simulated by presenting two slightly different images to each eye, motion parallax is provided by tracking the position and orientation of the user's head in order to update the rendered three-dimensional (3D) simulated scene in real time to reflect the change in viewpoint. Theoretically, there are many similarities between the information provided by these two depth cues [22]. A critical difference, however, is that the information obtained from motion parallax is heavily dependent on the contingency between a user's active movements and their resulting sensory consequences, which is not the case for stereopsis. The concept of sensorimotor contingencies [20, 21] has been repeatedly used in the presence literature, and it has been hypothesized that valid sensorimotor contingencies supported by the VR systems play a major role in establishing SOP in users [27].

Received Sept. 29, 2019; accepted for publication June 3, 2020; published online July 21, 2020. Associate Editor: Pierre Dragicevic.

2575-8144/2020/3(2)/020502/10/$00.00

1.1 *Related Work*

1.1.1 *Stereopsis and Motion Parallax*

A number of studies have investigated the relative importance of motion parallax and stereopsis for different behavioral tasks. The majority of these studies involve path tracing experiments where the participants' task was to find the correct endpoint of a line by tracking it from a highlighted starting point. The target line was presented among other distracting lines on a 3D display that allowed experiencing both motion parallax and stereopsis [19, 33]. Some studies applied other versions of a path tracing task, where participants were instructed to respond to whether two specified nodes presented in a complex interconnected graph were connected or not [10, 31, 34–36]. The greater importance of motion parallax over stereopsis for accurate judgments was shown in six of the path tracing studies [19, 31, 33–36]. Aygar and colleagues [2] tested the contribution of stereopsis and motion parallax by asking participants to respond to whether they could detect a specific target cluster in an artificially generated background pattern presented on a screen with shutter glasses and head tracking. When the frame rate of motion parallax (update rate of the display) was low, both motion parallax and stereopsis were needed in order to perform the task with high accuracy. However, the importance of stereopsis diminished when the update rate of motion parallax was increased. In most of the studies mentioned above, the combination of motion parallax and stereopsis led to the best performance. However, when only one or the other depth cue was provided for participants, the role of motion parallax dominated stereopsis.

However, there are also cases where stereopsis appeared to be more beneficial than motion parallax to accurately perform a task. Among the path tracing tasks, the only study that showed a benefit of stereopsis over motion parallax was reported by Hassaine and colleagues [10]. Furthermore, stereopsis seems to be more valuable in visual reaching tasks for objects. For instance, when participants had to move a ring along a wire curve in 3D space, Lion [14] found stereopsis to be more important than motion parallax. In another study, Bortiz and Booth [5] investigated the effect of motion parallax and stereopsis in a point location task in 3D space. Participants were required to move the tip of a pointer to a fixed point in a 3D environment. The results showed an improved performance under the stereopsis condition although introducing motion parallax had no effect on the results. Arsenault and Ware [1] obtained similar results and confirmed the importance of stereopsis in a comparatively similar reaching task as Bortiz and Booth [5]. In another behavioral study, Boustila and colleagues [6] implemented a paradigm to investigate the influence of motion parallax and stereopsis in the specific context of visiting houses in VR. They used a cave VR system to place participants in virtual houses and instructed them to verbally answer questions about the geometry of the rooms as well as their perceived difficulty of the task. The results indicated a main effect of stereopsis for judgments of geometry of the rooms. The perceived task difficulty was also higher when stereopsis was removed compared to the condition without motion parallax. They also evaluated participants' subjective feeling of presence using questionnaires. However, no significant difference was observed among the experimental conditions.

There are also other situations where the presence or absence of motion parallax and stereopsis seems to have no effect on the accuracy of a behavioral task. Barfield and Hendrix [3] conducted an experiment on a screen with shutter glasses and head tracking, where participants were instructed to match a virtual bent wire to three drawings printed on paper. No significant effect of motion parallax and stereopsis was reported. However, introducing motion parallax significantly increased participants' SOP as assessed by a presence questionnaire. Similar results in terms of SOP were also reported by Ijsselsteijn and colleagues [12].

In summary, the results of previous studies investigating the relative importance of motion parallax and stereopsis are controversial and inconsistent. Although a direct comparison across different studies is difficult due to methodological differences, several alternative yet compelling reasons might explain these inconsistencies. First, the relative importance of the two depth cues might differ depending on the nature of the task. For instance, motion parallax seems to be more important in path tracing tasks; however, stereopsis seems to be a more valuable cue in visually guided reaching tasks. Second, the complexity of the task might be a crucial factor. For example, Naepflin and Menozzi [19] showed that motion parallax is a stronger depth cue than stereopsis in more difficult tasks. According to Ware and Mitchel [37], the expertise of the participants is also important as motion parallax and stereopsis contribute equally to performance in more difficult tasks for novice participants. However, motion parallax is more helpful and dominates performance for expert participants.

1.1.2 *Sense of Presence*

The notion of presence has been around for a long time and measuring it properly in virtual environments has been a subject of considerable debate [4, 7, 8, 11, 13, 15, 16, 23–27, 29, 30, 32, 38]. Different questionnaires were introduced as a subjective measure of presence [29, 30, 38]. Presence questionnaires are useful to capture participants' conscious feelings and thoughts; however, they fail to detect unconscious and physiological reactions of participants. To address this issue, some physiological metrics such as heart rate, electrodermal activity (EDA), and skin temperature are commonly used as objective measures of presence in emotionally arousing virtual environments [17]. One interesting experimental scenario to which such measures were applied is the "pit room" experiment. The very early version of this paradigm was implemented by Gibson and Walk [9], where they used a visual cliff to test infants' perception of brinks and edges of cliffs. Inspired by the idea of the visual cliff, Slater and colleagues [30] introduced their own version of the "pit room" in VR. Meehan and colleagues [17] also used this paradigm to test the credibility of three physiological measures (heart rate, skin conductance, and skin temperature) as objective metrics for presence. In their version, participants wore a head-mounted display (HMD) VR system and were exposed to an intense situation in VR, which would be very frightening if it happened in reality: They were instructed to stand on a ledge of a very deep pit and drop objects in specified locations in the pit room. Even though participants were completely aware of the simulated nature of the experiment, their fear responses were significantly higher as compared to the non-threatening virtual room. The authors concluded that the changes in heart rate and EDA satisfied their requirements to objectively measure presence in virtual environments. A similar HMD-based paradigm was also used by Slater and colleagues [28], where they tested the effects of visual realism and the quality of VR renderings on the SOP.

1.2 *Study Aim*

The current study was designed to investigate the relative contribution of stereopsis and motion parallax to the SOP in two virtual environments using both subjective and objective measures. We used relatively similar paradigms as introduced in the previous literature [17, 28]. We focused on motion parallax and stereopsis because they are among the most prominent depth cues that provide similar information in terms of shape and depth. However, according to the theory of sensorimotor contingency [20, 21], they play very different roles, which determine their level of importance in establishing the SOP [27]. Considering the theory of sensorimotor contingency and Slater's notion of "place illusion," we expected motion parallax to be the main depth cue involved in creating the SOP in users.

In Experiment 1, we aimed for a paradigm that provides a means of measuring the SOP not only explicitly with a questionnaire but also with physiological measures and therefore more objective means. We designed a modified version of the pit room in an HMD-based virtual environment where participants' task was to cross a deep abyss on a narrow plank connecting two small platforms. Each participant completed the task under three experimental conditions: (1) when the lateral component of motion parallax was disabled, (2) when stereopsis was compromised by presenting the same image to both eyes in the HMD, (3) when the two depth cues were available to participants as in normal VR experience. We recorded participants' EDA during the experiment as an objective measure of presence. In addition, after each condition, we assessed participants' subjective SOP using the presence questionnaire introduced by Slater and colleagues [28]. The results from Experiment 1 showed that when motion parallax was restricted, the EDA responses were significantly higher compared to when stereopsis was disabled or when both stereopsis and motion parallax were enabled. Questionnaire scores, however, did not reveal any differences in subjective SOP between the conditions. We suspected that the frightening nature of the virtual environment and a high level of stress experienced in all three conditions might have overwhelmed participants and prevented them from consciously registering differences between the experimental conditions. To address this, in Experiment 2, we replicated the experimental conditions of Experiment 1 in a similar, but less frightening, virtual environment. We expected to get similar results as in Experiment 1 in terms of the EDA, and that these results would also be reflected in the questionnaire scores.

2. EXPERIMENT I

2.1 *Methods*

2.1.1 *Participants*

Twenty-six participants (14 females and 12 males) with an age between 19 and 30 years (mean 24.83 $\pm$ SD 3.6 years) were recruited for this study. All participants had normal or corrected-to-normal vision. Two of the participants withdrew from the study due to the frightening nature of the experiment. The statistical analysis was performed on the remaining 24 participants. The experiment was approved by the Human Ethics Committee of Queen's University, and all methods were in accordance with the Declaration of Helsinki. Participants were compensated with $5 for their participation.

2.1.2 *Apparatus and Stimuli*

The virtual environment was designed in Unity game engine and presented by means of an HTC VIVE HMD with a resolution of 1080 $\times$ 1200 pixels per eye and a refresh rate of 90 fps. An MSI backpack computer with Intel Core i7 CPU, 16 GB of RAM and an NVIDIA 1070 GTX graphic card was used for stimulus generation and control.

The virtual environment was modeled after the real laboratory space. The room was 12 $\times$ 10 m large; it contained a support pillar on one side and windows with black curtains on two sides. Both real and virtual spaces also contained two small platforms (5 cm high), which were placed flat on the ground. They were connected by a heavy wooden plank, which was 2.5 m long, 28 cm wide, and 5 cm thick. Upon entering the virtual environment, participants saw the same layout of the room with the same objects and furniture. It also included the two platforms and the connecting plank (Figure 1a, b). Note that the edges of the real plank and its ability to slightly bend between the two support platforms provided realistic tactile and haptic feedback.

Initially, there was no "pit" in the room. Later into the experiment, the floor of the virtual room was dropped by 8 m while the platforms and the plank remained at their locations, now high above the dropped floor (Fig. 1c).

We used a virtual humanoid avatar to represent the participant's body (Fig. 1d), with the intention to increase immersiveness in VR[1]. Participants could see "their" arms and lower body and particularly their feet. Seeing where they placed their feet was also thought to make them more confident and allowed them to move more naturally on the plank. Their movements were captured by tracking 12 reflective markers (three on each hand and foot) using 15 real-time Qualisys motion capture (MOCAP) cameras at a sampling rate of 120 Hz. We used the Final IK (Rootmotion Inc.) inverse kinematics plugin for the Unity game engine to predict the movement of other joints of the participants' body.

[1] For a video of the virtual environments and avatar renderings of Experiments 1 and 2, see www.youtube.com/channel/UCXsow3lmBmzTSaL9PILX0RQ/playlist

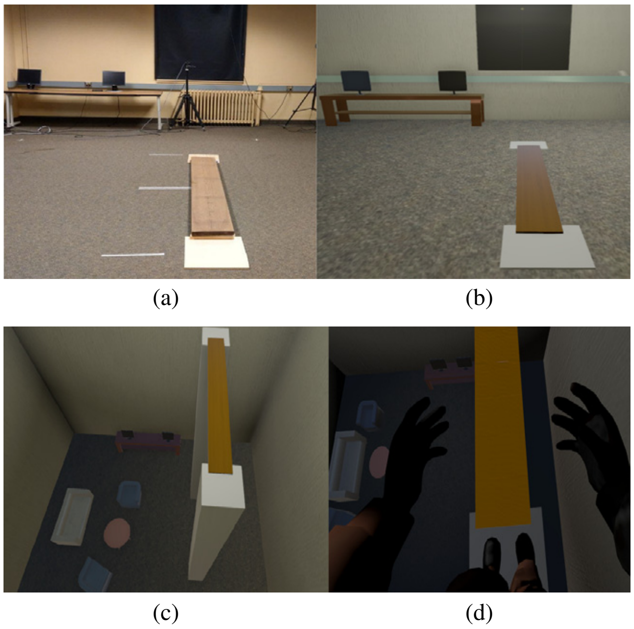

Figure 1. (a) Picture of the real environment and (b) screenshot of the virtual environment. (c) Screenshot of the pit room environment after the floor was dropped by 8 m. (d) First-person view of the humanoid avatar.

Participants' EDA was recorded using an E4 wristband (Empatica Inc.) at a sampling frequency of 4 Hz.

2.1.3 *Design*

The experiment employed a within-subject design in which each participant experienced three conditions. The experiment was run in one single session, which lasted approximately half an hour. The order of the three main conditions was counterbalanced across participants. The conditions were as follows:

(1) Participants experienced the pit room situation under standard VR conditions, that is, with both stereopsis and motion parallax functioning normally.

(2) Motion parallax was manipulated in the following way: While changes in angular orientation as well as displacements in posterior–anterior and vertical directions resulted in image changes that simulated normal motion parallax, lateral movement of the participants' head did not affect the rendering on the displays of the HMD. The reason behind disabling the lateral component of motion parallax was to make it comparable to stereopsis, which works only laterally. Stereopsis was rendered as in condition (1).

(3) Stereopsis was compromised. Rather than providing dichoptic images to the two eyes, both eyes were presented with the same image taken from a single "cyclopean" camera viewpoint centered between the eyes. Motion parallax was implemented as in (1).

2.1.4 *Task and Procedure*

The task and experimental procedure were explained to each participant before starting the experiment. Participants were unaware of the experimental manipulations; however, they were told that after a minute in the virtual environment, they would face a pit. Before the experiment started, participants were informed that they were free to withdraw from the experiment whenever they felt uncomfortable. Participants were then equipped with the E4 wristband and the HMD device, and the motion capture markers were attached to their feet and hands. When participants entered the virtual environment, they were instructed to stand at the center of the first platform. Participants were instructed to relax and look around in order to get familiar with the virtual environment while their baseline EDA was recorded. After 1 min, the experimenter dropped the virtual floor by pressing a button on the keyboard. The virtual floor started to drop smoothly by 8 m. Then, participants were asked to start walking on the plank, cross the abyss to reach the other platform, turn around, and come back to the starting position. Once participants reached their original position, the recording of EDA was stopped, and they took off the HMD. After the exposure, we immediately administered the presence questionnaire. Upon completion, we gave participants 3 min to rest. During this time, we asked participants to verbally report how difficult they perceived the experimental condition. After this relaxation period, participants put on the HMD again and experienced another condition, followed by the same presence questionnaire as in the previous condition and the verbal report on the task difficulty. This procedure was repeated once more for the third condition.

2.1.5 *Measurements*

We used two dependent measures: participants' EDA and subjective ratings of presence. The questionnaire consisted of 11 questions related to the experience of presence in the pit room [28]. Participants responded to each question on a seven-point Likert scale, where the numbers they chose reflected the level of agreement with the corresponding statement (1—not at all; 7—very).

The EDA was recorded for 1 min before the floor of the pit room was dropped. Subsequently, the EDA was recorded during the whole time the participants spent on the plank spanning the pit (2 min and 50 s on average). From the raw data, we derived ΔEDA as the difference between the measurements taken during the pit room condition and the baseline:

$$\Delta\text{EDA} = \text{mean EDA}_{\text{Pit room}} - \text{mean EDA}_{\text{Baseline}}. \quad (1)$$

The baseline EDA was recorded before each condition, and each condition was compared to its own baseline.

2.2 *Results*

Data from 24 participants were analyzed. Dropping the floor of the pit room had a significant effect on participants' EDA. We found a considerable increase in skin conductance in all three conditions (normal: $t(23) = -3.19$, $p = 0.004$; no-stereo: $t(23) = -3.5$, $p < 0.001$; no-parallax: $t(23) = -4.65$, $p < 0.001$).

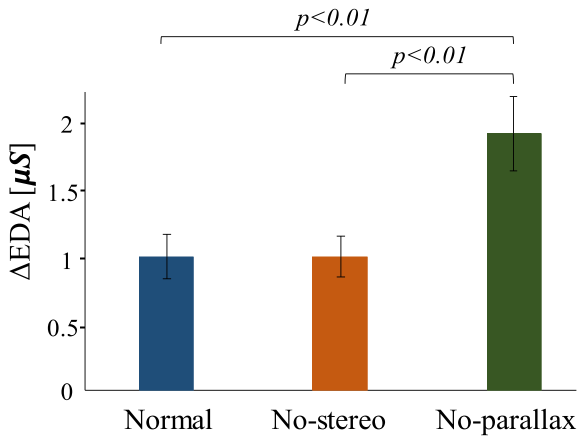

Figure 2. ΔEDA for normal, no-stereo, and no-parallax conditions. Error bars represent standard error of the mean.

Figure 2 illustrates the differences between the three different conditions. A two-way repeated measures analysis of variance (ANOVA) (condition × order) revealed a main effect of condition $F(2, 54) = 6.71$, $p = 0.002$. Bonferroni-corrected t-tests with an adjusted alpha level of 0.017 per test (0.05/3) showed a significant difference between the condition without motion parallax and the two other conditions. However, no significant difference was observed between the normal and no-stereo conditions ($p = 0.99$). Further, the verbal reports from the participants indicated that the majority of them felt more unstable and perceived the walking task in the no-parallax condition as more difficult than in the other two conditions (19 out of 24 participants).

We neither observed a significant main effect of the order of exposure ($F(5, 54) = 2.34$, $p = 0.054$) nor was the interaction of condition by order significant ($F(10, 54) = 0.45$, $p = 0.91$).

We also calculated the mean score of the questionnaires for each condition (Figure 3). A two-factor (condition × order) repeated measures ANOVA on the mean questionnaire scores revealed no significant difference in terms of subjective feeling of presence across the conditions, $F(2, 54) = 0.04$, $p = 0.95$. The interaction between condition and order was also not significant ($p = 0.06$).

These results are not in line with our expectation that compromising either stereopsis or motion parallax would lead to lower subjective SOP as compared to the normal condition. Visual inspection of the means of the individual participants also did not imply that specific effects on particular questions canceled each other out in the summary statistics.

The time it took for participants to complete the experiment was similar across the different conditions (normal: $M = 183.5$ s, SD = 16.06, range: 132–212; no-stereo: $M = 186.87$ s, SD = 28.48, range: 117–260; no-parallax: $M = 191.08$ s, SD = 23.47, range: 132–245).

2.3 *Discussion*

Data from the first experiment revealed that participants' EDA was significantly increased in the condition with

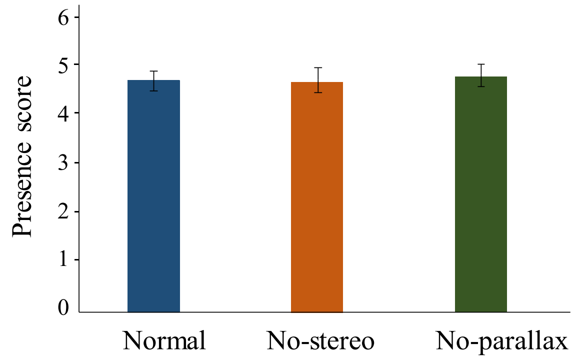

Figure 3. Mean presence scores for normal, no-stereo, and no-parallax conditions. Error bars represent standard error of the mean.

restricted motion parallax compared to the other two conditions. However, we were unable to detect the same effect on EDA responses by eliminating stereopsis when compared to the normal condition. We expected that eliminating an essential depth cue such as motion parallax affects the EDA response in the opposite way: If EDA is considered an objective measure of presence and if we expect that compromising motion parallax breaks the place illusion, then a lower EDA response would be expected in these conditions. A possible explanation for why this was not the case in Experiment 1 is that a fearful environment like a pit room becomes less naturalistic but scarier and more stressful when an important depth cue is missing. In our experiment, participants found themselves in a potentially dangerous situation in which their health seemed to depend on staying upright and balanced on the narrow plank. The situation might have been perceived even more dangerous in the no-parallax condition because participants were then lacking an important sensory cue that helped them to maintain their balance. Consequently, EDA increases significantly. Although these results question the reliability of EDA as an objective measure of presence (see General Discussion for more details), they still support our expectation of a larger effect for motion parallax as compared to stereopsis in our pit room experiment. Observing the fact that EDA responses were significantly higher in the situation where we impaired motion parallax indicates the importance of this depth cue in the pit room paradigm.

The results from the presence questionnaire were surprisingly flat across the conditions. This contradicts our hypothesis that compromising an important depth cue such as motion parallax would affect presence scores in a similar way as previously shown in other studies [3, 12]. We can only speculate why we were unable to replicate those results. First, VR is still a new experience for the majority of people including for most of our participants. It is possible that the new high-quality VR experience was so exciting to them that it prevented them from consciously differentiating among the conditions. Second, the presence questionnaire might not be sensitive enough to capture differences in SOP across the three conditions of this experiment. Third, we suspected that the nature of our pit room was so scary and overwhelming for the participants that they mainly focused their attention on crossing the plank safely without falling down. Thus, they might not have been able to consciously notice the differences between the conditions (even though most of them felt the task was more difficult to complete in the no-parallax condition). To test whether the frightening nature of the experiment affected participants' experienced presence in the virtual environment, we conducted a second experiment using a less scary version of the pit room paradigm.

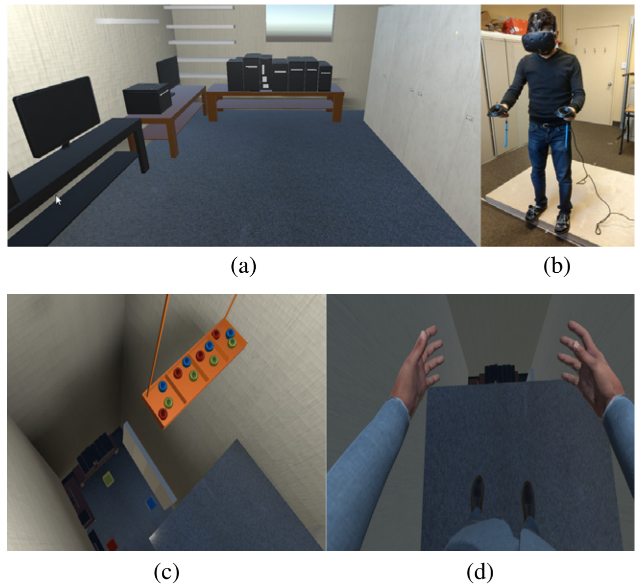

Figure 4. (a) Virtual room that was viewed by the participant while standing on a wooden platform in the real room (b). (c) Pit room environment after the floor was lowered by 8 m. (d) First-person view of the humanoid avatar.

3. EXPERIMENT 2

3.1 *Methods*

3.1.1 *Participants*

Eighteen participants (9 females and 9 males) between 20 and 32 years of age (mean 25.55 $\pm$ SD 3.14 years) were recruited for this study. All participants had normal or corrected-to-normal vision. The experiment was approved by the Human Ethics Committee of Queen's University and all methods were in accordance with the Declaration of Helsinki. Participants were compensated with $5 for their participation.

3.1.2 *Apparatus and Stimuli*

The experimental set-up, apparatus, and measurement methods were similar to Experiment 1. However, we used a different virtual room, replaced the plank with a ledge, and employed a slightly different task. The no-stereo condition was also implemented differently (see below).

The virtual room was again modeled after the real room in which the experiment took place. This room was smaller than that used in Experiment 1. It was 5 m high, 3.5 m long, and 2.5 m wide, and it contained one window, shelves on the wall, and some older computers alongside with a big gray closet on one side. All these features were implemented in the virtual room (Figure 4a). We placed a wooden platform, which was 120 cm long, 100 cm wide, and 14 cm thick in the real room in order to provide realistic tactile and haptic feedback (Fig. 4b). Like in Experiment 1, there was no pit in the beginning. After the baseline period, a part of the virtual floor was dropped by 8 m. The remaining part of the virtual floor matched the area of the wooden platform in the real room. Participants were now standing at the edge of a ledge, looking down into the abyss. At the same time, a small tablet appeared in front of the participant containing 12 balls with different colors (Fig. 4c). A virtual humanoid self-avatar was used for all participants (Fig. 4d). The height of the avatar was adjusted to the height of each participant. Two VIVE controllers and two VIVE trackers were used to track participants' hands and feet, respectively. Participants were able to see their lower body and arms in VR. The Final IK (Rootmotion Inc.) inverse kinematic plugin was used to simulate the motion of other body parts using the information derived from participants' feet and hands (controllers and VIVE trackers).

3.1.3 *Design*

The experiment consisted of three different conditions. Similar to Experiment 1, participants experienced the virtual environment with normal motion parallax and stereopsis, with impaired motion parallax but normal stereopsis, or without stereopsis but normal motion parallax. For compromising stereopsis, we decided to introduce a monocular viewing condition because we suspected that presenting identical images to both eyes, as done in Experiment 1, is not the best way of eliminating stereopsis. Although the method used in Experiment 1 eliminates relative disparity over the stimulus, convergence between the eyes still indicates a distance at infinity. Monocular viewing was achieved by presenting a dark screen to participants' non-dominant eye while the image presented to the dominant eye remained active.

A within-subject design was implemented with the order of conditions being counterbalanced across participants. The experiment took approximately 30 min in a single session.

3.1.4 *Task and Procedure*

We explained the procedure of the experiment in full detail to each participant before starting the session. Similar to Experiment 1, participants were unaware of the experimental manipulations. However, they were informed that they would be seeing a pit. Participants were free to withdraw from the experiment if they felt discomfort at any stage of the experiment. Before starting the experiment, each participant's dominant eye was determined using the Miles ocular dominance test [18]. Participants' feet were tracked by two VIVE trackers mounted on a pair of sandals. Participants were instructed to hold the VIVE controllers, which were used to represent their virtual hands. The height of the avatar was adjusted to each participant's self-reported height. Afterward, they were equipped with the E4 wristband and the HMD to start the experiment. We asked participants to look around and familiarize themselves with the virtual environment for 1 min while we started recording the baseline EDA response. After 1 min, the virtual floor smoothly moved down by 8 m. The task for participants was to grab the virtual balls in front of them with their virtual hands and throw them into baskets located on the floor of the pit room. Participants were told to throw each ball into the basket that matched the color of the ball. Introducing this task aimed at forcing participants to look down into the space of the lowered floor. Participants were familiarized with the use of the controllers and the virtual hands to grab and throw the balls in one practice trial before starting the experiment. In the practice trial, they experienced the normal virtual room without the lowered floor.

After completing the task, participants took off the HMD and immediately started to fill out the presence questionnaire. Upon completion, we gave participants a 3-minute break to relax between the conditions. Similar to Experiment 1, we verbally asked for participants' opinion on the difficulty of the task during the break. After the relaxation period, they put on the HMD again and completed the same task in a different experimental condition followed by the same presence questionnaire. This procedure was repeated once more for the last condition. The order of the three conditions was counterbalanced between participants.

3.1.5 *Measurements*

As in Experiment 1, we used the presence questionnaire [28] and participants' EDA as our dependent measures and Eq. (1) to derive ΔEDA.

3.2 *Results*

The data of all 18 participants were included in the analysis. Figure 5 summarizes the results for ΔEDA. Like in Experiment 1, we found that the EDA increased in all experimental conditions as compared to the baseline. As expected under the less scary environment of Experiment 2, the increase was less pronounced than in Experiment 1 (0.86 μS instead of 1.31 μS, on average).

Fig. 5 illustrates the results of EDA responses for the three conditions. A significant difference among conditions was revealed by a two-way repeated measures ANOVA (condition × order), $F(2, 36) = 8.19$, $p = 0.001$. Bonferroni-corrected t-tests with an adjusted alpha level of 0.017 per test (0.05/3) revealed a significant difference between the condition without motion parallax and the other two conditions. The absence of motion parallax led to a significant increase in the EDA. The normal and no-stereo conditions were not significantly different ($p = 0.11$).

We found no effect for the order of exposure $F(5, 36) = 0.94$, $p = 0.46$. There was also no interaction between experimental conditions and the order of exposure $F(10, 36) = 0.84$, $p = 0.59$. Similar to Experiment 1, most of the participants reported that they perceived it more difficult to complete the task and stay balanced on the wooden ledge in the no-parallax condition (17 out of 18 participants).

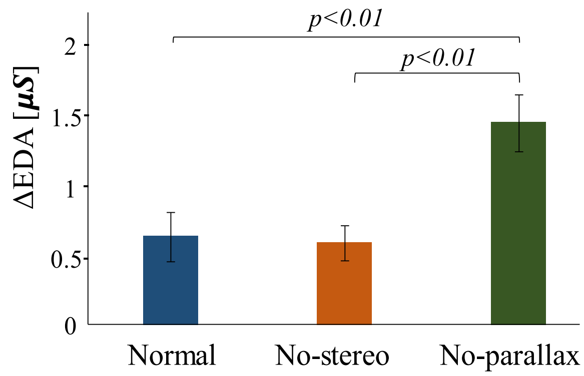

Figure 5. ΔEDA for normal, no-stereo, and no-parallax conditions. Error bars represent standard error of the mean.

Mean scores of the presence questionnaire were also calculated across 11 questions for all three conditions. A two-way repeated measures ANOVA (condition × order) revealed a significant main effect of condition $F(2, 36) = 4.87$, $p = 0.013$ (Figure 6). Bonferroni-corrected t-tests with an adjusted alpha level of 0.017 per test (0.05/3) showed a significant difference between the no-parallax and normal conditions, $t(17) = 2.73$, $p = 0.014$. The no-stereo condition did not differ significantly from the normal condition ($p = 0.3$) and the no-parallax condition ($p = 0.021$).

The time it took participants to complete the experiment was similar across the different conditions (normal: $M = 239.83$ s, SD = 33.65, range: 171–278; no-stereo: $M = 232.05$ s, SD = 39.9, range: 181–302; no-parallax: $M = 246.44$ s, SD = 28.60, range: 196–301).

3.3 *Discussion*

Similar to Experiment 1, the EDA responses differed among conditions. Participants' EDA responses were significantly higher in the no-parallax condition compared to the other two conditions. Eliminating stereopsis did not have the same effect on the EDA response, and it remained at the same level as in the normal condition.

Contrary to Experiment 1, the results from the presence questionnaire showed that restricting motion parallax affected participants' subjective SOP. The mean presence score for the condition without motion parallax was significantly lower compared to the normal condition. The mean presence score in the condition without stereopsis also appears to be slightly lower compared to the normal condition; this difference, however, did not reach significance. The results suggest that impairing motion parallax affects subjective SOP more strongly than when impairing stereopsis.

It should be noted that we used a different way for eliminating stereopsis in this experiment. In Experiment 1, we presented identical images to both eyes of the participants, while we used monocular viewing in Experiment 2. We reasoned that the monocular viewing might be a better

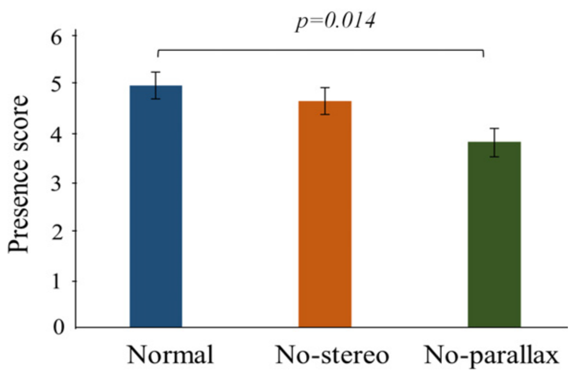

Figure 6. Mean presence scores for normal, no-stereo, and no-parallax conditions. Error bars represent standard error of the mean.

option to eliminate stereopsis, and if there was a difference between the two options, then the results of the no-stereo condition in Experiment 2 might also be different from those in Experiment 1. This was, however, not the case. We observed no difference in the pattern of the EDA results and the questionnaire scores between the non-stereo condition and the normal viewing condition of Experiments 1 and 2. In both experiments, the results of the non-stereo viewing condition were not statistically different from those of the normal condition.

4. GENERAL DISCUSSION

In two experiments, we investigated the relative contribution of motion parallax and stereopsis to the SOP in VR using two versions of the pit room paradigm. We tested the effects of these two depth cues on both objective and subjective measures that the previous literature has established for the assessment of the SOP in virtual environments [17, 28]. For an objective measure of SOP, we recorded participants' EDA. In order to measure participants' subjective feeling of presence, we adopted the presence questionnaire developed by Slater and colleagues [28]. We hypothesized that compromising one of the depth cues would lead to reduced SOP and that this would be reflected in both measures. Assuming an important role of sensorimotor contingencies in establishing SOP in the virtual environment [27], we further expected a stronger effect when compromising motion parallax as compared to stereopsis.

With respect to the results of the presence questionnaire, we did not observe any difference across the conditions in Experiment 1. We speculated that the virtual environment and the participants' task to walk across the plank were so stressful and overwhelming that it prevented participants from paying attention to differences between the experimental conditions. Indeed, when exposing participants to a less scary environment in Experiment 2, the non-parallax condition resulted in lower subjective SOP as compared to the normal condition. This supports our interpretation of the results of the presence scores in Experiment 1. As we expected, Experiment 2 allowed participants to consciously monitor the virtual environment and respond to differences between conditions.

The EDA results show that when motion parallax was compromised, participants' EDA responses were significantly higher than in the condition without stereopsis and when both depth cues were provided. Our interpretation of the finding is the following. The reason why EDA responses increased in the no-parallax condition is probably not related to the decreased SOP. Rather, the absence of motion parallax deprives the participant of important sensory feedback to keep upright—not just in VR but also in the real world of the laboratory. Although the HMD still responds naturally to movement in the posterior–anterior direction, it does not provide the expected sensory update in response to lateral movements. That in itself might be very upsetting, and it is probably even more upsetting in a situation that implies a deep pit and the threat of falling into it. Although the direction of the difference between the no-parallax condition and the other conditions was not aligned with our expectations, it did show that eliminating motion parallax made the virtual environment more scary compared to the normal condition.

Our results and the above interpretation suggest that the EDA is not necessarily a good proxy to measure SOP in virtual environments. Although it has been argued that a greater SOP will evoke greater affective and related physiological responses in virtual environments [17], similar reactions could be elicited in situations where the reliability of the visual cues relevant to the task performance is reduced. In Experiment 1, eliminating motion parallax resulted in strongly enhanced EDA. However, subjective SOP did not change at all in the no-parallax condition compared to the other two conditions. The same pattern was also shown in Experiment 2, where we observed that the less realistic, yet more scary situation (no-parallax condition) resulted in strongly enhanced EDA, while the subjective SOP decreased significantly in comparison to the normal condition. Thus, a greater physiological response does not necessarily mean a higher SOP even in a stressful virtual environment.

It is important to mention that in this study, we only eliminated the lateral component of motion parallax, while still allowing for parallactic information resulting from head rotation, vertical movements, and movements along the posterior–anterior direction. Stereopsis as a depth cue provides very similar information as motion parallax to the visual system; however, stereopsis works only in the lateral direction. Thus, in order to make the conditions more comparable, we only eliminated the lateral information of motion parallax. Further, some parallactic information was necessary for participants in order to complete the required tasks. For instance, participants needed head rotation and information along the posterior–anterior and vertical directions when they moved forward on the plank or when they leaned over the wooden ledge to see the baskets located at the base of the abyss.

One limitation of this study is that we did not test for participants' stereo-vision. However, when examining the individual data, we did not observe any obvious outliers in the results that would result from a participant being stereo-blind. Also note that failing to identify stereo-blindness in a participant would have weakened observed effects. It is unlikely that they could have resulted in rejecting the null hypothesis based on spurious effects.

5. CONCLUSION AND FUTURE DIRECTIONS

In summary, this study provides insight into the role of two important depth cues, motion parallax and stereopsis, in the SOP in VR. The results of the presence questionnaire in Experiment 2 indicated that impaired motion parallax led to a significantly lower SOP as compared to impaired stereopsis. Furthermore, both experiments showed that when motion parallax is eliminated, participants' EDA was significantly increased. Although open questions remain about the relation between physiological reactions and perceived presence in VR, our experiments provide an example in which increased EDA is obviously not coupled with an increased SOP.

In these two experiments, we only disabled the lateral component of motion parallax due to the experimental design. Future studies and different experimental designs are needed to investigate the effect of eliminating different components of motion parallax on the SOP. Furthermore, participants' stereo-vision should be examined using standard tests either in VR or existing ones (e.g., Randot stereotest) in future studies investigating the relative importance of motion parallax and stereopsis.

In this study, we targeted the EDA as one possible objective measure of the SOP. However, the EDA is not the only physiological measure that has been proposed as an objective measure of presence. An interesting future study that could further enrich the literature of the SOP is to investigate the reliability of other physiological and behavioral measures such as ECG, EEG, respiration, skin temperature, and performance in a similar context.

REFERENCES

[1] R. Arsenault and C. Ware, "The importance of stereo and eye-coupled perspective for eye-hand coordination in fish tank VR," Presence Teleoperators Virtual Environ. **13**, 549 (2004).

[2] E. Aygar, C. Ware, and D. Rogers, "The contribution of stereoscopic and motion depth cues to the perception of structures in 3D point clouds," ACM Trans. Appl. Percept **15**, 1 (2018).

[3] W. Barfield and C. Hendrix, "Visualizing the structure of virtual objects using head tracked stereoscopic displays," *Proc. of IEEE 1997 Annual Int'l. Symposium on Virtual Reality* (IEEE, Piscataway, NJ, 1997), p. 114.

[4] W. Barfield, d Zeltzer, T. Sheridan, and M. Slater, "Presence and performance within virtual environments," *Virtual Environ. Adv. Interface Des.* (Oxford University Press, Inc.198 Madison Ave. New York, NY, 1995), 475.

[5] J. Boritz and K. S. Booth, "A study of interactive 3D point location in a computer simulated virtual environment," *Proc. ACM Symposium on Virtual Reality Software and Technology* (Association for Computing Machinery, New York, NY, 1997), p. 181.

[6] S. Boustila, D. Bechmann, and A. Capobianco, "Effects of adding visual cues on distance estimation, presence and simulator sickness during virtual visits using wall screen," *Proc. Computer Graphics International Conference on - CGI '17* (Association for Computing Machinery, New York, NY, 2017), p. 21.

[7] S. R. Ellis, "Presence of mind: a reaction to Thomas Sheridan's 'further musings on the psychophysics of presence'," Presence Teleoperators Virtual Environ. **5**, 247 (1996).

[8] J. Freeman, S. E. Avons, D. Pearson, D. Harrison, and N. Lodge, "Behavioral realism as a metric of presence," *1st Int. Wkshp. on Presence* (International Society for Presence Research, 1998).

[9] E. J. Gibson and R. D. Walk, "The 'visual cliff'," Sci. Am. **202**, 64–71 (1960).

[10] D. Hassaine, N. S. Holliman, and S. P. Liversedge, "Investigating the performance of path-searching tasks in depth on multiview displays," ACM Trans. Appl. Percept **8**, 8 (2010).

[11] W. A. IJsselsteijn and H. De Ridder, "Measuring temporal variations in presence," Presence Shar. Virtual Environ. Work. Univ. Coll. Lond. **10** (1998).

[12] W. Ijsselsteijn, H. De Ridder, J. Freeman, S. E. Avons, and D. Bouwhuis, "Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence," Presence Teleoperators Virtual Environ. **10**, 298 (2001).

[13] K. M. Lee, "Presence, explicated," Commun. Theory **14**, 27 (2004).

[14] D. M. Lion, "Three dimensional manual tracking using a head-tracked stereoscopic display," Doctoral dissertation, University of Washington, (1993).

[15] M. Lombard and T. Ditton, "At the heart of it all: the concept of presence," J. Comput. Commun. **3**, JCMC321 (1997).

[16] R. McCall, S. O'Neil, and F. Carroll, "Measuring presence in virtual environments," *Conf. on Human Factors in Computing Systems: CHI'04 Extended Abstracts on Human Factors in Computing Systems* (Association for Computing Machinery, Special Interest Group on ComputerHuman Interaction New York, NY, 2004), p. 783.

[17] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks, "Physiological measures of presence in stressful virtual environments," ACM Trans. Graph. **21**, 645 (2002).

[18] W. R. Miles, "Ocular dominance in human adults,," J. Gen. Psychol. **3**, 412 (1930).

[19] U. Naepflin and M. Menozzi, "Can movement parallax compensate lacking stereopsis in spatial explorative search tasks?," Displays **22**, 157 (2001).

[20] A. Noë, *Action in Perception* (MIT Press, Cambridge, Massachusetts, 2004).

[21] J. K. O. Regan and A. Noë, "A sensorimotor account of vision and visual consciousness," Behav. Brain Sci. **24**, 939 (2001).

[22] B. Rogers and M. Graham, "Similarities between motion parallax and stereopsis in human depth perception," Vis. Res. **22**, 261 (1982).

[23] H. T. Regenbrecht and T. W. Schubert, "Measuring presence in virtual environments," *Proc. Human Computer Interface Int'l., San Francisco* (Springer, New York, 1997).

[24] T. Schubert, F. Firedmann, and H. Regenbrecht, "Embodied presence in virtual environments," *Visual Representations and Interpretations* (Springer-Verlag, London, 1999), p. 269.

[25] T. B. Sheridan, "Further musings on the psychophysics of presence," Presence Teleoperators Virtual Environ. **5**, 241 (1996).

[26] M. Slater, "Measuring presence: a response to the Witmer and Singer presence questionnaire," Presence **8**, 560 (1999).

[27] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," Phil. Trans. R. Soc. B Biol. Sci. **364**, 3557 (2009).

[28] M. Slater, P. Khanna, J. Mortensen, and I. Yu, "Visual realism enhances realistic response in an immersive virtual environment," IEEE Comput. Graph. Appl. **29**, 76 (2009).

[29] M. Slater, M. Usoh, and A. Steed, "Depth of presence in virtual environments," Presence Teleoperators Virtual Environ. **3**, 130 (1994).

[30] M. Slater, M. Usoh, and A. Steed, "Taking steps: the influence of a walking technique on presence in virtual reality," ACM Trans. Comput. Interact. **2**, 201 (1995).

[31] R. L. Sollenberger and P. Milgram, "Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task," Hum. Factors **35**, 483 (1993).

[32] M. Usoh, K. W. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr, "Walking > walking-in-place > flying," *Virtual Environments, Proc. 26th Annual Conf. on Computer Graphics and Int eractive Techniques* (ACM Press, New York, NY, 1999), p. 359.

[33] M. H. P. H. Van Beurden, A. Kuijsters, and W. A. IJsselsteijn, "Performance of a path tracing task using stereoscopic and motion based depth cues," *2010 2nd Int'l. Workshop on Quality of Multimedia Experience, QoMEX 2010 -Proceedings* (IEEE, Piscataway, NJ, 2010), p. 176.

[34] C. Ware, K. Arthur, and K. S. Booth, "Fish tank virtual reality," *Proc. INTERACT'93 and CHI'93 Conf. on Human Factors in Computing Systems* (ACM, New York, NY, 1993), p. 37.

[35] C. Ware and G. Franck, "Evaluating stereo and motion cues for visualizing information nets in three dimensions," ACM Trans. Graph. **15**, 121 (1996).

[36] C. Ware and P. Mitchell, "Reevaluating stereo and motion cues for visualizing graphs in three dimensions," *APGV '05 Proc. 2nd Symp. Appl. Percept. Graph. Vis* (Association for Computing Machinery, New York, NY, 2005), p. 51.

[37] C. Ware and P. Mitchell, "Visualizing graphs in three dimensions," ACM Trans. Appl. Percept **5**, 2 (2008).

[38] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: a presence questionnaire," Presence **7**, 225 (1998).