GAMES FOR HEALTH JOURNAL: Research, Development, and Clinical Applications
Volume 9, Number 4, 2020

DOI: 10.1089/g4h.2019.0102

Studying the Effect of Display Type and Viewing Perspective on User Experience in Virtual Reality Exergames

Wenge Xu, MSc,[1] Hai-Ning Liang, PhD,[1] Zeying Zhang, BSc,[1] and Nilufar Baghaei, PhD[2,3]

Abstract

Background: Physical inactivity has been identified as the fourth leading cause of death globally. It is now well established that a sedentary lifestyle is a unique risk factor for several diseases such as type 2 diabetes and cardiovascular disease, which account for about 30% of global mortality. Diabetes is a major preventable cause of costly and debilitating renal failure, heart disease, lower limb amputation, and avoidable blindness. In recent years, the idea of using interactive computing systems that leverage gamification to promote physical activity has been widely researched. Prior studies have shown that exergames, that is those that encourage physical activity, can increase enjoyment and intrinsic motivation compared with conventional exercises; as such, they can be effective in promoting physical and mental health. There has been some research on immersive virtual reality (VR) exergames; however, to the best of our knowledge, it is limited and preliminary. This work aims at filling the gap and investigates the effect of display type (DT) and viewing perspective (VP) on players' exertion, engagement, and overall game experience in immersive VR exergames.
Objective: This article aims at examining whether DT and VP can affect gameplay performance, players' exertion, game experience, cybersickness, and electroencephalography (EEG) engagement index when playing a gesture-based (i.e., body motion) exergame.
Materials and Methods: Study 1 employed a one-way between-subjects design with 24 participants equally distributed in two groups (immersive VR and 50-inch TV) to perform 12 pre-defined gestures. The main outcome measures were National Aeronautics and Space Administration-Task Load Index (NASA-TLX) workload for each group as well as 7 Likert scale and EEG engagement index for each gesture. Study 2 included 16 participants in playing a game with the gestures selected from study 1. All participants played 4 versions based on combinations of DT (immersive VR and 50-inch TV) and VP (first-person and third-person) to assess exertion (%HR_{max}, calories consumption, and Borg RPE 6–20), game experience, cybersickness, and EEG engagement index.
Results: Study 1 results showed that DT had no effect on the ratings of the gestures, NASA-TLX workload, and EEG engagement index. Study 2 results showed that immersive VR not only resulted in a significantly higher exertion (%HR_{max}, calories consumption, and Borg RPE) but also helped achieve better positive game experience in challenge, flow, sensory and imaginative immersion, as well as lower negative affect. We also found that nausea and oculomotor were significantly higher in immersive VR.
Conclusion: This pilot study demonstrates that youth who played gesture-based exergame in immersive VR had a higher level of exertion (%$HR_{max,}$ calories consumption, and Borg RPE), although the number of performed gestures were not significantly different. They also felt that immersive VR was much more challenging, immersive (flow, sensory and imaginative immersion), and had a lower negative affect than a 50-inch TV; however, immersive VR was more likely to make youth have higher cybersickness.

Keywords: Virtual reality, Exergame, Motion-based gaming, Head-mounted display, Large display, Viewing perspective

[1]Department of Computer Science and Software Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, China.
[2]Department of Information Technology, Otago Polytechnic Auckland International Campus, Auckland, New Zealand.
[3]School of Natural and Computational Sciences, Massey University Auckland, Auckland, New Zealand.

Introduction

Physical inactivity has been identified as the fourth leading cause of death globally.[1] It is now well established that a sedentary lifestyle is a unique risk factor for several diseases such as type 2 diabetes and cardiovascular disease,[2] which account for about 30% of global mortality. In recent years, the idea of using interactive computing systems that leverage gamification to promote physical activity has been widely researched.[3] Prior studies[4–7] have shown that exergames, a type of games that encourage physical activity, can increase enjoyment and intrinsic motivation compared with conventional exercises; as such, they can be effective in promoting physical and mental health.[8,9]

Given the advantages of engaging people in long-term and regular physical activity, various non-immersive virtual reality (VR)[10] (like using interfaces such as a flat-screen TV/monitor) exergames have been designed to encourage people to be more active,[11] promote a positive lifestyle[12] and self-care.[13] Previous literature has shown that exergames could bring physical and health outcomes to players. For example, Peng et al.[14] have performed a meta-analysis of energy expenditure in exergames where their main finding suggests that exergames are as effective as traditional physical activities that facilitate light- and moderate-intensity physical exertion. Huang et al.[15] found that exergames can induce positive changes in happiness, perceived energy levels, and relaxation for people who are enthusiastic about doing exercises. Other studies have shown that exergames are as effective as conventional balance training exercises.[16,17] Moreover, the benefits of playing exergames include, but not limited to, improving the quality of life,[18] reducing state anxiety,[19] as well as improvements in the number of steps taken, standing balance, gait speed, and mobility.[20]

Given the recent emergence of immersive VR technology,[10] which frequently uses head-mounted displays (HMDs), there is limited and only preliminary research on immersive VR exergames. Recently, Barathi et al.[21] have implemented an exercycle game with interactive feedforward by using immersive VR to improve players' performance and maintain intrinsic motivation. Ioannou et al.[22] found that virtual augmented running and jumping in immersive VR could increase intrinsic motivation, perceived competence, and flow. Xu et al.[23] have found that playing exergame in immersive VR would not result in a higher cybersickness than a 50-inch TV. In general, researchers have suggested that immersive VR is useful in promoting physical activity in sedentary and obese children,[24] especially to increase their motivation to exercise.[25,26] However, the difference between exergaming with a common display and immersive VR is still largely underexplored, especially regarding their physical and health benefits.

Traditional approaches such as direct observations[27] and subjective measurements[11] are the commonly used methods to measure user experience during games. However, they can be intrusive and not reliable. Psychophysiological methods, such as using electroencephalography (EEG), provide relatively non-intrusive, covert, and reliable measurements of affective states that determine user experience, and this makes them suitable for studying interactive entertainment.[28] Such methods have been used to investigate the effect of controller types,[29] viewing angles,[30] display types (DTs), and tasking modes[23] on players' brainwave patterns.

Chang et al.[31] and Stoffregen et al.[32] have proved that videogames can carry a significant risk of cybersickness. One solution to reduce it is by seeking the most suitable viewing perspective (VP) (e.g., first person vs. third person). For example, Medina et al.[33] found that cybersickness was more pronounced for the first-person viewing perspective (1PP) group than the third person viewing perspective (3PP) group when performing locomotion walking in navigation tasks in an immersive VR environment. Similarly, Monteiro et al.[34] pointed out that playing an immersive VR racing game in 3PP is less likely to induce cybersickness when compared with playing it in 1PP.

Given the considerations just mentioned, the aim of this study was to investigate the effect of DT (immersive VR and large TV) and VP (1PP and 3PP) on players' exertion, engagement, and overall gaming experience of exergames. To this end, we conducted a first study to select a gesture set for a gesture-based game to make sure that the selected gestures would not affect players' gameplay in both DTs. Afterward, in a second study, we investigated the effect of DTs and VPs when interacting with an exergame.

The current investigation has been guided by the following hypotheses. Because previous research[7] showed that playing an exercycle game with a common flat monitor and immersive VR led to an equal level of burned calories, we hypothesized that:

H1: **(a)** There would be no significant differences in gameplay performance (i.e., completing the same number of gestures) among DTs; therefore, **(b)** we believe the levels of exertion (%HR_{max}, calories burned, and Borg RPE) should also be the same among the DTs.

H2: Immersive VR could result in a higher game experience than large display (LD).

Similarly, because prior work that tested different types of interventions showed that 3PP could lead to a lower motion sickness than 1PP,[33,34] we predicted that:

H3: During a gameplay of more than 3 minutes, **(a)** 3PP could lead to a lower cybersickness than 1PP in exergames. As for immersive VR, we believe that **(b)** it could lead to a higher level of cybersickness than LD.

Materials and Methods

Study 1

In the interest of removing any possible bias toward a DT, Study 1 aimed at identifying a set of full-body gestures for the exergame to be used in Study 2. That is, we evaluated gestures that would not be affected by DT.

Participants. Twenty-four participants were recruited from a local university campus to participate in this experiment. Because two participants' EEG data were lost due to bad connection between the devices, we recruited another two participants. The final 24 participants (6 females) were aged between 19 and 27 (mean = 22.04) years old. Twenty-two played videogames regularly (17 of them played weekly). For the immersive VR group, only two of them were frequent users of immersive VR.

The inclusion criteria of the participants for the study were those who: (1) answered ''no'' to all Physical Activity

Readiness Questionnaires,[35] (2) had a resting blood pressure lower than 140/90 mmHg, and (3) had a common (10%–90%)[36] resting heart rate depending on their age and gender.

Instruments. To avoid familiarity with gestures that could potentially affect the selection of gestures, we employed a one-way between-subjects experiment design with 24 participants (6 females) equally distributed in two groups where the independent variable was DT—immersive VR and LD. The experiment was conducted at a university lab. We used an Oculus Rift CV1 as our VR HMD and a 50-inch 4K TV as our LD. Both devices were connected to a standard computer with an i7 CPU, 16GB RAM, and a GeForce GTX 1080Ti GPU. The brainwave signals were collected by a MUSE headset Edition 1. The program was built in Unity3D, and players' gestures were detected by a Microsoft Kinect 2.

The National Aeronautics and Space Administration-Task Load Index (NASA-TLX)[37] is a validated instrument for measuring workload,[38] which consists of six subscales that represent independent clusters of variables: mental, physical, and temporal demands, frustration, effort, and performance. It first presents users with a series of pairs of rating scale titles (e.g., effort vs. mental demands) and asks users to choose which of the items was more important to the experience of workload in the task(s) that were just performed. Then, it asks users to rate each workload cluster in a 21-Likert scale.[38] The NASA-TLX has been widely used by universities, industries, and governments.[39]

Participants' rating. Participants needed to rate each gesture via a 7-point Likert scale, ranging from 1, strongly disagree, to 7, strongly agree. A higher score indicated that participants would like to have such a gesture in the final version of the game.

The EEG metric we used for this study was the engagement index, which has been widely used in the research of biocybernetics and automation systems,[40–44] and is a measurement of how cognitively engaged a person is in a task.[40] It can be calculated by the formula $E = \frac{\beta}{(\alpha+\theta)}$[41] where α, β, and θ are averaged values of α, β, and θ waves from the EEG device (i.e., MUSE 1).

Task: performing the gestures. Participants needed to perform 12 different gestures in a computer program (Fig. 1), which was developed by the researchers, with the TV or immersive VR device depending on their assigned group. All gestures were evaluated by rehabilitation doctors we had access to. There were six simple gestures (*Psi*: raising two hands; *Squat*: performing a squat; *Kick*: raising any leg; *Walk*: performing walk-in-place; *Wheel*: performing steering wheel motion; *Zoom*: leaning arms forward and stretching them out), and six complex gestures that were combinations of simple gestures (*Squat+Psi*; *Squat+Wheel*; *Kick+Zoom*; *Kick+Wheel*; *Walk+Psi*; and *Walk+Zoom*). For each gesture, instructions were given to participants via a pre-recorded 5-second video (Fig. 1a). Then, they were requested to repeat each gesture in two 10-second sessions, with 5 seconds of rest in between. The order of the gestures was counterbalanced during the experiment.

Procedure. Before the experiment, participants were told about the purpose of the experiment, given the information

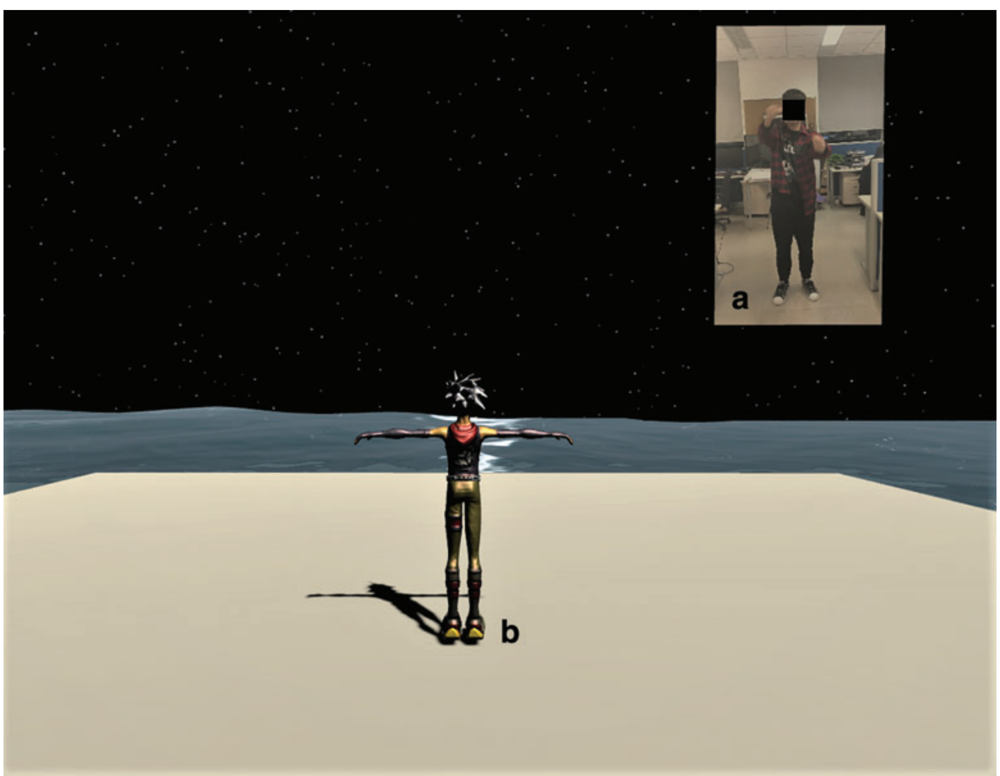

FIG. 1. Screenshot of Study 1 program. **(a)** Video display area for participants to follow. **(b)** A character represents the participant. Color images are available online.

Table 1. Means (Standard Deviations) of National Aeronautics and Space Administration-Task Load Index Questionnaire Results

	Overall	*Mental*	*Physical*	*Temporal*	*Performance*	*Effort*	*Frustration*
VR	49.11 (14.67)	40.42 (19.82)	52.92 (19.71)	48.75 (19.08)	45.83 (19.75)	43.33 (18.26)	25.83 (20.10)
LD	47.42 (9.15)	34.17 (19.29)	61.25 (16.53)	43.75 (9.08)	39.17 (20.87)	41.67 (9.85)	17.50 (12.70)

LD, large display; VR, virtual reality.

sheet to read, and the consent form to sign. Once they agreed to participate, participants were asked to complete a pre-experiment questionnaire to collect demographic data. After the devices used in the experiment were described to them, a researcher helped calibrate the MUSE (to make sure that the MUSE had a good connection with the MUSE application running on a mobile device).

After they understood the process, participants proceeded to play the computer program and perform the gestures. After the experiment, participants needed to complete the post-experiment questionnaire and give comments on the gestures to the experimenter through an interview. The whole experiment lasted about 40 minutes for each participant. The experiment was conducted under the supervision of the experimenter, and the surroundings were cleared of any obstacles to give a safe environment to the participants.

Statistical analysis. SPSS version 24 for windows was used for analysis. The Kolmogorov–Smirnov test was used to verify the normality of the data. For NASA-TLX[37] overall workload, we analyzed the data by using a univariate analysis of variance (ANOVA) and for its subscales, we employed a multivariate ANOVA to evaluate the effects of DT on gestures that had been performed by participants. For participants' ratings and the EEG engagement index, we employed a mix-design ANOVA with gesture (12 gestures) as the within-subjects variable and DT as the between-subjects variable. Bonferroni correction was used for pairwise comparisons, and Greenhouse-Geisser adjustment was used for degrees of freedom if there were violations to sphericity in the data.

Results

NASA-Task Load Index. A univariate ANOVA yielded no significant effect of DT ($F_{1,22} = 0.115$, $P = 0.737$) on overall workload. A multivariate ANOVA also showed no significant effect of DT on the six NASA-TLX subscales: mental ($P = 0.442$), physical ($P = 0.274$), temporal ($P = 0.421$), performance ($P = 0.430$), effort ($P = 0.783$), and frustration ($P = 0.283$). See Table 1 for results.

Gesture set

Participants' ratings. Results of participants' ratings of each gesture can be found in Table 2. ANOVA tests yielded a significant effect of gesture ($F_{5.539,121.848} = 4.288$, $P < 0.001$) but not of gesture × group ($F_{11,242} = 0.970$, $P = 0.474$) on the rating scores of the gestures. There was no significant effect of group ($F_{1,22} = 0.049$, $P = 0.826$) on participants' rating of each gesture. *Post hoc* pairwise comparisons revealed significant differences between gesture *Psi – Kick+Zoom*, *Psi – Kick+Wheel*, *Walk – Kick+Zoom* (all $P < 0.05$).

EEG engagement index. ANOVA tests yielded no significant effect of gesture ($F_{11,242} = 1.727$, $P = 0.175$), group ($F_{1,22} = 2.619$, $P = 0.120$), or gesture × group ($F_{11,242} = 0.712$, $P = 0.726$) on task engagement for each gesture. Results of the EEG engagement index of each gesture can be found in Table 2.

Discussion

Our results indicated that DT did not affect players' preference of the gestures, their workload, and the engagement

Table 2. Means (Standard Deviations) of Participants' Ratings and Electroencephalography Engagement Index Results of Each Gesture

	Participants' ratings		*EEG engagement index*	
Gesture	*VR*	*LD*	*VR*	*LD*
Psi	5.92 (0.79)	5.67 (1.16)	0.81 (0.48)	0.51 (0.44)
Squat	4.92 (1.38)	4.92 (1.98)	0.76 (0.33)	0.64 (0.27)
Kick	5.08 (1.08)	5.58 (1.38)	0.74 (0.37)	0.52 (0.41)
Walk	5.83 (0.84)	5.67 (1.07)	1.06 (1.48)	0.71 (0.81)
Wheel	5.00 (1.28)	5.25 (1.82)	0.77 (0.37)	0.55 (0.23)
Zoom	5.58 (1.17)	5.17 (1.47)	0.95 (0.75)	0.47 (0.35)
Squat+Psi	5.42 (1.17)	4.67 (1.88)	0.82 (0.83)	0.55 (0.27)
Squat+Wheel	4.83 (1.59)	4.17 (1.70)	0.66 (0.44)	0.48 (0.36)
Kick+Zoom	4.08 (1.51)	5.00 (1.13)	0.65 (0.44)	0.36 (0.32)
Kick+Wheel	4.17 (1.40)	4.25 (1.42)	0.70 (0.58)	0.54 (0.24)
Walk+Psi	5.50 (1.57)	5.25 (1.29)	0.25 (1.01)	0.46 (0.18)
Walk+Zoom	5.50 (1.57)	5.33 (1.37)	0.59 (0.38)	0.46 (0.23)

EEG, electroencephalography.

FIG. 2. The six blocks that were used in the game: (left to right) Kick, Squat, Zoom, Psi, Squat+Psi, and Kick+Zoom. Color images are available online.

index when performing these full-body gestures. We also observed that some gestures might raise issues for future gameplay. We therefore selected the gesture set with the following exclusion considerations: (1) Based on the participants' ratings and comments, we decided to exclude *Wheel*, *Squat+Wheel*, and *Kick+Wheel* gestures since the ratings of these gestures were low. In addition, 20 out of 24 participants complained during the interview that performing these gestures was too hard (e.g., P9 from the non-VR group: ''This gesture is too difficult to do''). (2) Based on our observations, we decided to exclude *Walk*, *Walk+Psi*, and *Walk+Zoom* gestures since participants could easily go forward instead of walking-in-place when performing such gestures, which could cause tracking issues because, similar to nearly all motion tracking devices, the Kinect 2 we used in Study 2 only had a limited operational tracking area.

In summary, our exergame in the second study was designed to have four simple gestures—*Psi*, *Squat*, *Kick*, *Zoom*, and two complex gestures—*Squat+Psi* and *Kick+Zoom*.

Since task engagement index was the same, therefore we hypothesize that:

H4: DT and VP would not affect the EEG task engagement index.

Study 2

In Study 2, we investigated the impact of DT (Large TV and Immersive VR) and VP (1PP and 3PP) on gesture-based exergame gameplay performance and experience.

Participants. Another 16 participants were recruited for this study. Because one participant's EEG data were lost due

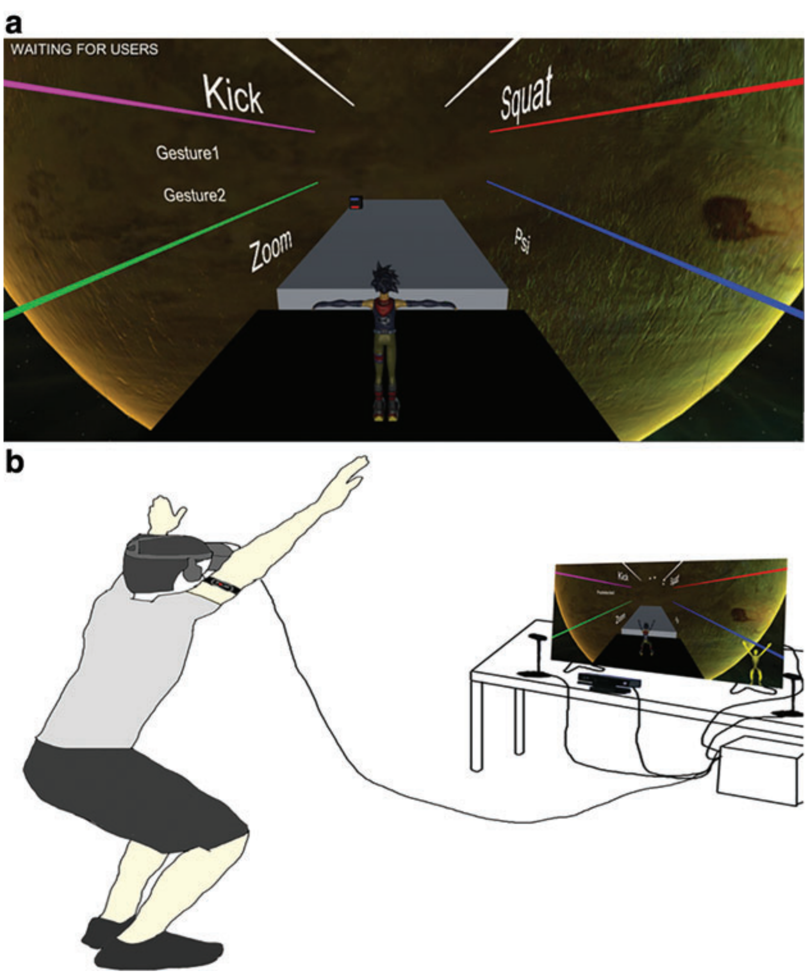

FIG. 3. A screenshot of the exergame where the name and colored lines in the game work as a reminder of each gesture for the player (**a**; top) and an example of a participant performing Psi+Squat gesture during the game (**b**; bottom). Color images are available online.

TABLE 3. *P* VALUES OF TWO-WAY REPEATED ANALYSIS OF VARIANCE RESULTS ON GAME PERFORMANCE, EXERTION, SIMULATOR SICKNESS QUESTIONNAIRE, AND ELECTROENCEPHALOGRAPHY TASK INDEX

	Game performance	*%HR$_{max}$*	*Calories burned*	*Borg RPE*	*Nausea*	*Oculomotor*	*EEG engagement index*
DT	0.468	<0.01**	<0.01**	<0.001***	<0.05*	<0.01**	0.439
VP	0.338	0.852	0.320	0.403	0.072	0.812	0.446
DT × VP	0.929	0.086	1.000	0.333	0.300	0.510	0.303

Level of significance: *<0.05, ** <0.01, and *** <0.001.
DT, display type; VP, viewing perspective.

to bad connection, we recruited one more participant. The final 16 participants (5 females) included in the data analysis were between the ages of 18 and 28 years (mean = 21.75). Ten of them had some prior experience with immersive VR (2 of them interacted with it weekly). Fifteen participants played videogames regularly (12 of them weekly).

We used the same inclusion criteria as Study 1 for this study.

Instruments. The experiment followed a 2 × 2 within-subjects design with combinations of (1) VP—(1PP and 3PP) and (2) DT—(immersive VR and LD). The order of DT × VP was counterbalanced in the experiment.

In addition to the devices used in Study 1, we used a Polar OH1, which has been proved to be able to capture good HR data when compared with the gold standard of HR measurement of an electrocardiography device,[45,46] to record participants' heart rate and calorie consumption.

Participants' task performance was evaluated in terms of the percentage of blocks removed (i.e., when the gesture was performed correctly).

Participants' game experience was measured by using the 33-item core module of the Game Experience Questionnaire.[47] It consists of seven components: competence, sensory and imaginative immersion, flow, tension, challenge, negative affect, and positive affect.

Cybersickness was assessed by using the 16-item Simulator Sickness Questionnaire.[48] It measures a wide range of possible symptoms of cybersickness, including (but not limited to) nausea, eyestrain, dizziness, and vertigo. Each symptom was rated on a severity scale that ranged from 0 (none) to 3 (severe). The scale had an observed Cronbach's α of 0.91. This scale was aggregated to produce 2 measures of cybersickness (Nausea and Oculomotor) with 27 and 21 points, respectively.

Exertion was evaluated by (1) the average heart rate (%HR$_{max}$) and was expressed as a percentage of a participant's estimated maximum HR (220 minus age).[49] (2) Calories burned and (3) ratings of perceived exertion were measured by the Borg RPE 6–20 scale.[50]

Physiological involvement was assessed by the EEG engagement index. For details of this measurement, see Study 1, Instruments section.

Participants' preference of the conditions (VR-1PP, VR-3PP, LD-1PP, and LD-3PP) was measured by their rankings of the condition from 1 to 4, where 1 stood for the most preferred option and 4 for the least preferred option.

Task: GestureStar game. Inspired by the commercial exergames *Beat Saber* and *Just Dance*, we developed GestureStar. In GestureStar, players encountered blocks flying toward them every 6 seconds and were required to make the corresponding gesture to eliminate each block within 6 seconds; otherwise, they would miss it. One game lasted about 8 minutes (1 minute for training and 7 minutes for the actual experiment). In total, participants were required to perform 10 gestures during training and 70 during gameplay.

As stated earlier, the game had four simple gestures (*Psi*, *Squat*, *Kick*, and *Zoom*), and two complex gestures (*Squat+Psi* and *Kick+Zoom*). We employed six different blocks to represent the gestures in the game (Fig. 2). Figure 3a shows a screenshot of the game, and Figure 3b shows the setup of a player playing the game.

Procedure. Participants were briefed of the purpose of the experiment and asked to sign the consent form and complete a pre-experiment questionnaire. Afterward, a researcher helped participants to wear and calibrate the MUSE 1 and Polar OH1. We only recorded EEG and heart rate data for the 7-minute experimental part. After each condition, participants were asked to complete the post-experiment questionnaire. They could rest as much as they want between conditions. After the experiment, they were asked to give feedback and rank each condition. The whole experiment lasted about 1 hour for each participant.

Statistical analysis. Similar to Study 1, SPSS version 24 for windows was used for analysis. The Kolmogorov–Smirnov test was used to verify the normality of the data. We

TABLE 4. MEANS (STANDARD DEVIATIONS) OF COMPLETION RATES, EXERTION, NAUSEA, OCULOMOTOR, AND ELECTROENCEPHALOGRAPHY ENGAGEMENT INDEX

	Completion rate	*%HR$_{max}$*	*Calories burned*	*Borg RPE*	*Nausea*	*Oculomotor*	*EEG engagement index*
VR_1PP	91.79% (3.28%)	53.60% (6.82%)	42.81 (13.11)	14.50 (1.90)	1.88 (1.54)	2.81 (2.54)	0.36 (0.23)
VR_3PP	92.77% (3.36%)	52.78% (6.33%)	43.81 (13.38)	13.94 (1.12)	2.56 (2.13)	3.13 (2.22)	0.26 (0.43)
LD_1PP	92.41% (5.64%)	50.60% (6.03%)	34.75 (12.07)	11.94 (1.57)	1.69 (1.89)	1.69 (1.82)	0.35 (0.30)
LD_3PP	93.57% (4.92%)	51.25% (5.77%)	35.75 (13.00)	12.06 (1.84)	1.50 (1.41)	1.50 (1.41)	0.38 (0.22)

1PP, first-person viewing perspective; 3PP, third person viewing perspective.

Table 5. *P* Values of Two-Way Repeated Analysis of Variance Results of the Game Experience Questionnaire

	Competence	*Sensory and imaginative immersion*	*Flow*	*Tension*	*Challenge*	*Negative affect*	*Positive affect*
DT	0.588	<0.01**	<0.01**	0.730	<0.01**	<0.05*	0.125
VP	0.181	0.284	0.070	0.453	0.133	0.060	0.348
DT × VP	0.085	<0.01**	0.073	0.224	0.118	0.770	<0.05*

Level of significance: *<0.05, ** <0.01, and *** <0.001.

used the two-way repeated measures ANOVA and Bonferroni correction for pairwise comparisons.

Results

Hypothesis testing

Analytical results of game performance, exertion (average $\%HR_{max}$, Calories burned, and Borg RPE), simulator sickness questionnaire, and EEG engagement index can be found in Table 3.

Details of participants' task performance and exertion for each condition can be found in Table 4. No significance was found on task performance between conditions, supporting **H1a**. However, immersive VR had led to a higher $\%HR_{max}$ ($P=0.005$), calories burned ($P=0.001$), and Borg RPE rating ($P=0.000$) than LD, not supporting **H1b**.

Analytical results of each Game Experience Questionnaire component are shown in Table 5. The score for immersive VR was higher than LD regarding challenge ($P=0.002$), flow ($P=0.004$), and sensory and imaginative immersion ($P=0.002$); whereas immersive VR had a lower score regarding negative affect ($P=0.023$) than LD. Therefore, the results supported the **H2**. Table 6 shows the scores for each component.

No significance was found for Nausea and Oculomotor on VP, not supporting **H3a**. **H3b** was supported since immersive VR had caused a higher level of Nausea ($P=0.016$) and Oculomotor ($P=0.010$) than LD. Details of the sickness scores for each condition can be found in Table 4.

H4 was supported as no significant effect of DT and VP was found on EEG engagement index. Values of EEG engagement index can be found in Table 4.

User preference

Friedman tests yielded a significant difference depending on which version participants preferred $\chi^2(3)=10.059$, $P=0.018$. However, *post hoc* analysis with Wilcoxon signed-rank tests and Bonferroni correction did not reveal any significant difference between conditions, although 63% of the participants selected VR-1PP as their top choice.

Discussion

Discussion on the hypotheses

We found support in our results for **H1a**, where participants completed the same number of gestures in both immersive VR and LD conditions. However, **H1b** was not supported, even though the completion rates of the gestures were the same. One possible explanation might be because the weight of the VR HMD that participants had to carry during the immersive VR condition increased the intensity of the exergame, although the Oculus CV1 just weighted 470 g.

We found support for **H2**; that is, playing exergames in immersive VR had a better gameplay experience as it was more challenging, immersive (based on the flow, sensory, and imaginative immersion components) to participants, and had less negative effects. Interestingly, our findings did not support the results from a previous study[23] in which researchers found that playing a motion-based exergame in immersive VR might have the same level of game experience. One possible explanation might be because the length of our game was much longer than theirs.

Previous studies[33,34] suggested that 3PP could lead to a lower sickness level than 1PP; however, we did not find support for **H3a**. That is, playing an exergame in 3PP did not result in a lower cybersickness level than in 1PP. We hypothesize that since our game demanded a reasonable amount of movement, the bone vibration equated in lower levels of cybersickness in both versions equally.[51] Further, in our experiment, participants often focused on a fixed point, so they could better observe the oncoming objects, which equated to the same advantage as 3PP, thus not bringing any special advantage in this scenario. **H3b** was supported, as our data indicated that players felt sicker (both nausea and oculomotor) when playing in immersive VR than LD, which is in line with previous immersive VR studies.[52,53]

We confirmed our **H4** that DT and VP did not affect the EEG engagement index.

Practical implications

Our results indicate that playing a full-body gesture exergame in immersive VR could lead to a higher exertion

Table 6. Means (Standard Deviations) of Game Experience Questionnaire Subscales

	Competence	*Sensory and imaginative immersion*	*Flow*	*Tension*	*Challenge*	*Negative affect*	*Positive affect*
VR_1PP	2.79 (0.58)	2.66 (0.52)	2.51 (0.69)	0.90 (0.74)	2.09 (0.85)	0.86 (0.84)	3.09 (0.83)
VR_3PP	2.74 (0.69)	2.38 (0.63)	2.26 (0.70)	0.98 (0.75)	2.06 (0.79)	1.10 (0.81)	2.78 (1.03)
LD_1PP	2.54 (0.55)	2.01 (0.51)	1.78 (0.46)	1.00 (0.82)	1.81 (0.85)	1.19 (0.78)	2.64 (0.75)
LD_3PP	2.88 (0.63)	2.09 (0.61)	1.74 (0.43)	0.81 (0.78)	1.50 (0.56)	1.34 (0.76)	2.78 (0.72)

level than LD (i.e., it burned more calories, and led to a higher $\%HR_{max}$, and perceived exertion level on the Borg RPE). Moreover, playing an exergame in immersive VR can induce not only a higher immersion level but also a lower negative feeling than LD. As such, when players need some exercise, they could be introduced to playing exergames with VR HMDs. However, if players start to get cybersick quickly, they should play exergames with LD.

For game designers, consideration should be taken with respect to gestures: (1) by not designing and including complex gestures (e.g., wheel used in Study 1); (2) by avoiding gestures such as walk-in-place because players might need to move around, which could lead to tracking issues and potentially dangerous situations.

Strengths, limitations, and future work

The strengths of our research include: (1) the gestures used for the exergame were selected systematically (from Study 1) to remove any bias toward any particular type of display that could originate from a gesture; (2) the effect of DT (immersive VR and LD) and VP (1PP and 3PP) on cybersickness and exertion in exergames were never previously examined. To our knowledge, we are the first ones to conduct this research; (3) another strength of the study is that it has contributed to the limited research topic of immersive VR on health benefits to its users (e.g., exertion).

There are some limitations to this research. One limitation is that the research involved a relatively small sample (though this is normal in research published in this area).[54] Future work can involve a larger and more diverse group of participants. Moreover, the current version of GestureStar seems only to be a light-intensity game as participants' $\%HR_{max}$ is lower than 64% (Table 4), which is the lower bound of moderate intensity exersises.[55] One possible solution to increase the intensity of the game is by narrowing the wait time for the next block if the player eliminates the current block in advance. In addition, future work can focus on reducing potential nausea and other adverse side effects while increasing the intensity of the immersive VR version of the exergame.

Acknowledgments

The authors thank all the participants who volunteered their time and the reviewers for their suggestions that helped them improve this article.

Ethical Approval and Consent to Participate

The study was approved by Research Ethics Sub-Committee of Xi'an Jiaotong-Liverpool University with code (EXT 19-01-01). Informed consent was obtained from all participants who volunteered to participate in the two studies.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This research was supported by XJTLU Key Program Special Fund (Grant No. KSF-A-03) and XJTLU Research Development Fund.

References

1. Lee I-M, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. Lancet 2012; 380:219–229.
2. Williams G, Fruhbeck G. *Obesity: Science to Practice*. Chichester, UK: Wiley; 2009.
3. Mandryk RL, Gerling KM, Stanley KG. Designing games to discourage sedentary behaviour. In: Nijholt A, ed. *Playful User Interfaces*. Singapore: Springer Singapore; 2014; pp. 253–274.
4. Bailey BW. Energy cost of exergaming: A comparison of the energy cost of 6 forms of exergaming. Arch Pediatr Adolesc Med 2011; 165:597.
5. Finkelstein S, Suma EA. Astrojumper: Motivating exercise with an immersive virtual reality exergame. Presence 2011; 20:78–92.
6. Monedero J, Lyons EJ, O'Gorman DJ. Interactive video game cycling leads to higher energy expenditure and is more enjoyable than conventional exercise in adults. PLoS One 2015; 10:e0118470.
7. Shaw LA, Marks S, Wunsche BC, et al. Development and evaluation of an exercycle game using immersive technologies. In: Maeder A, Warren J, eds. *Australasian Workshop on Health Informatics and Knowledge Management (HIKM)*. Vol. 164. Sydney, Australia: Australian Computer Society; 2015; pp. 75–85.
8. Pluchino A, Lee SY, Asfour S, et al. Pilot study comparing changes in postural control after training using a video game balance board program and 2 standard activity-based balance intervention programs. Arch Phys Med Rehab 2012; 93:1138–1146.
9. Rosenberg D, Depp CA, Vahia IV, et al. Exergames for subsyndromal depression in older adults: A pilot study of a novel intervention. Am J Geriatr Psychiatry 2010; 18:221–226.
10. Zeng N, Pope Z, Lee J, Gao Z. Virtual reality exercise for anxiety and depression: A preliminary review of current research in an emerging field. J Clin Med 2018; 7:42.
11. Gerling K, Livingston I, Nacke L, Mandryk R. Full-body motion-based game interaction for older adults. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems—CHI'12*. Austin, TX: ACM Press; 2012; p. 1873.
12. Gao Y, Mandryk RL. GrabApple: The design of a casual exergame. In: Anacleto JC, Fels S, Graham N, et al., eds. *Entertainment Computing—ICEC 2011*. Vol. 6972. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011; pp. 35–46.
13. Faiola A, Kharrazi H. Diabetes education and serious gaming: Teaching adolescents to cope with diabetes. In: Hayes BM, Aspray W, eds. *Health Informatics: A Patient Centered Approach to Diabetes*. Cambridge, MA: MIT Press; 2010; pp. 151–177.
14. Peng W, Lin J-HT, Crouse WJ. Is playing exergames really exercising? A meta-analysis of energy expenditure in active video games. Cyberpsychol Behav Soc Netw 2011; 14: 681–688.
15. Huang H-C, Nguyen HV, Cheng TCE, et al. A randomized controlled trial on the role of enthusiasm about exergames: Players' perceptions of exercise. Games Health J 2019; 8: 220–226.
16. Sápi M, Domján A, Fehérné Kiss A, Pintér S. Is Kinect training superior to conventional balance training for

healthy older adults to improve postural control? Games Health J 2018; 8:41–48.
17. Bacha JMR, Gomes GCV, Freitas TB, et al. Effects of Kinect adventures games versus conventional physical therapy on postural control in elderly people: A randomized controlled trial. Games Health J 2017; 7:24–36.
18. Silva AR, Iunes DH, Carvalho JM, et al. Effects of exergaming on quality of life in cancer patients. Games Health J 2018; 7:385–392.
19. Viana RB, Alves CL, Vieira CA, et al. Anxiolytic effects of a single session of the Exergame Zumba® Fitness on healthy young women. Games Health J 2017; 6:365–370.
20. Garcia JA, Schoene D, Lord SR, et al. A bespoke Kinect stepping exergame for improving physical and cognitive function in older people: A pilot study. Games Health J 2016; 5:382–388.
21. Barathi SC, O'Neill E, Lutteroth C, et al. Interactive feedforward for improving performance and maintaining intrinsic motivation in VR exergaming. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI'18*. Montreal QC, Canada: ACM Press; 2018; pp. 1–14.
22. Ioannou C, Archard P, O'Neill E, Lutteroth C. Virtual performance augmentation in an immersive jump & run exergame. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*. Glasgow, Scotland, UK: ACM Press; 2019; pp. 1–15.
23. Xu W, Liang H-N, Yu Y, et al. Assessing the effects of a full-body motion-based exergame in virtual reality. In: *Proceedings of the Seventh International Symposium of Chinese CHI on—Chinese CHI 2019*. Xiamen, China: ACM Press; 2019.
24. Rizzo AS, Lange B, Suma EA, Bolas M. Virtual reality and interactive digital game technology: New tools to address obesity and diabetes. J Diabetes Sci Technol 2011; 5:256–264.
25. Mestre DR, Dagonneau V, Mercier C-S. Does virtual reality enhance exercise performance, enjoyment, and dissociation? An exploratory study on a stationary bike apparatus. Presence 2011; 20:1–14.
26. Plante TG, Aldridge A, Su D, et al. Does virtual reality enhance the management of stress when paired with exercise? An exploratory study. Int J Stress Manag 2003; 10: 203–216.
27. Norton J, Wingrave CA, LaViola JJ. Exploring strategies and guidelines for developing full body video game interfaces. In: *Proceedings of the Fifth International Conference on the Foundations of Digital Games—FDG'10*. Monterey, CA: ACM Press; 2010; pp. 155–162.
28. Ravaja N. Contributions of psychophysiology to media research: Review and recommendations. Media Psychol 2004; 6:193–235.
29. Nacke LE. Wiimote vs. controller: Electroencephalographic measurement of affective gameplay interaction. In: *Proceedings of the International Academic Conference on the Future of Game Design and Technology—Futureplay'10*. Vancouver, BC, Canada: ACM Press; 2010; p. 159.
30. Monteiro D, Liang H-N, Zhao Y, Abel A. Comparing event related arousal-valence and focus among different viewing perspectives in VR gaming. In: Ren J, Hussain A, Zheng J, et al., eds. *Advances in Brain Inspired Cognitive Systems*. Cham: Springer International Publishing; 2018; pp. 770–779.
31. Chang C-H, Pan W-W, Tseng L-Y, Stoffregen TA. Postural activity and motion sickness during video game play in children and adults. Exp Brain Res 2012; 217:299–309.
32. Stoffregen T, Faugloire E, Yoshida K, et al. Motion sickness and postural sway in console video games. Hum Factors 2008; 50:322–331.
33. Medina E, Fruland R, Weghorst S. Virtusphere: Walking in a human size VR ''Hamster Ball.'' Proc Hum Factors Ergon Soc Annu Meet 2008; 52:2102–2106.
34. Monteiro D, Liang H-N, Xu W, et al. Evaluating enjoyment, presence, and emulator sickness in VR games based on first- and third-person viewing perspectives. Comput Animat Virt Worlds 2018; 29:e1830.
35. Thomas S, Reading J, Shephard R. Revision of the Physical Activity Readness Questionnaire (PAR-Q). Can J Sport Sci 1993; 17:338–345.
36. Ostchega Y, Porter K, Hughes J, et al. Resting pulse rate reference data for children, adolescents, and adults: United States, 1999–2008. Natl Health Stat Rep 2011; 41: 1–16.
37. Hart SG. *NASA Task load Index (TLX). Volume 1.0; Paper and Pencil Package*. Moffett Field, CA: NASA Ames Research Center; 1986.
38. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv Psychol 1988; 52:139–183.
39. Hart SG. Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proc Hum Factors Ergon Soc Annu Meet 2006; 50: 904–908.
40. Freeman FG, Mikulka PJ, Prinzel LJ, Scerbo MW. Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. Biol Psychol 1999; 50: 61–76.
41. Pope AT, Bogart EH, Bartolome DS. Biocybernetic system evaluates indices of operator engagement in automated task. Biol Psychol 1995; 40:187–195.
42. Dorneich M, Whitlow S, Ververs P, et al. Closing the loop of an adaptive system with cognitive state. Proc Hum Factors Ergon Soc Annu Meet 2004; 48:590–594.
43. Mikulka PJ, Scerbo MW, Freeman FG. Effects of a biocybernetic system on vigilance performance. Hum Factors 2002; 44:654–664.
44. Chaouachi M, Frasson C. Exploring the relationship between learner EEG mental engagement and affect. In: Aleven V, Kay J, Mostow J, eds. *Intelligent Tutoring Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010; pp. 291–293.
45. Hettiarachchi IT, Hanoun S, Nahavandi D, Nahavandi S. Validation of Polar OH1 optical heart rate sensor for moderate and high intensity physical activities. PLoS One 2019; 14:e0217288.
46. Schubert M, Clark A, Rosa A. The Polar® OH1 optical heart rate sensor is valid during moderate-vigorous exercise. Sports Med Int Open 2018; 2:E67–E70.
47. IJsselsteijn WA, De Kort YAW, Poels K. *The Game Experience Questionnaire*. Eindhoven: Technische Universiteit Eindhoven; 2008.
48. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993; 3: 203–220.
49. Kluwer W. *ACSM's Guidelines for Exercise Testing and Prescription*. Philadelphia, PA: Lippincott Williams & Wilkins; 2013.

50. Borg G. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982; 14:377–381.
51. Weech S, Moon J, Troje NF. Influence of bone-conducted vibration on simulator sickness in virtual reality. PLoS One 2018; 13:e0194137.
52. Akiduki H, Nishiike S, Watanabe H, et al. Visual-vestibular conflict induced by virtual reality in humans. Neurosci Lett 2003; 340:197–200.
53. Sharples S, Cobb S, Moody A, Wilson JR. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. Displays 2008; 29:58–69.
54. Caine K. Local standards for sample size at CHI. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI'16*. Santa Clara, CA: ACM Press; 2016; pp. 981–992.
55. Canning KL, Brown RE, Jamnik VK, et al. Individuals underestimate moderate and vigorous intensity physical activity. PLoS One 2014; 9:e97927.

Address correspondence to:
Hai-Ning Liang, PhD
Department of Computer Science and Software Engineering
Xi'an Jiaotong-Liverpool University
Room SD433, 111 Ren'ai Road
Suzhou. Jiangsu 215123
China

E-mail: haining.liang@xjtlu.edu.cn