Physiological Effectivity and User Experience of Immersive Gait Rehabilitation

Negin Hamzeheinejad[a*] Daniel Roth[a†] Daniel Götz[a‡] Franz Weilbach[b§] Marc Erich Latoschik[a¶]

[a]Human-Computer Interaction, Julius-Maximilians-Universität Würzburg, Germany
[b]Klinik Bavaria, Bad Kissingen, Germany

Figure 1: Different scenes users visit during their stroll in the virtual environment while they perform the walking exercises. From left to right: grassland, forest, stream land, and beach.

Abstract

Gait impairments from neurological injuries require repeated and exhaustive physical exercises for rehabilitation. Prolonged physical training in clinical environments can easily become frustrating and de-motivating for various reasons which in turn risks to decrease efficiency during the healing process. This paper introduces an immersive VR system for gait rehabilitation which targets user experience and increase of motivation while evoking comparable physiological responses needed for successful training effects. The system provides a virtual environment consisting of open fields, forest, mountains, waterfalls, animals, and a beach for inspiring strolls and is able to include a virtual trainer as a companion during the walks. We evaluated the ecological validity of the system with healthy subjects before performing the clinical trial. We assessed the system's target qualities with a longitudinal study with 45 healthy participants in three consecutive days in comparison to a baseline non-VR condition. The system was able to evoke similar physiological responses. The workload was increased for the VR condition but the system also elicited a higher enjoyment and motivation which was the main goal. The latter benefits slightly decreased over time (as did workload) while they were still higher than in the non-VR condition. The virtual trainer did not show to be beneficial, the corresponding implications are discussed. Overall, the approach shows promising results which renders the system a viable alternative for the given use case while it motivates interesting direction for future work.

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

*e-mail: negin.hamzeheinejad@stud-mail.uni-wuerzburg.de
†e-mail: daniel.roth@uni-wuerzburg.de
‡e-mail: daniel.goetz1@stud-mail.uni-wuerzburg.de
§e-mail: weilbach.f@klinik-bavaria.com
¶e-mail: marc.latoschik@uni-wuerzburg.de

1 Introduction

Traumatic neural injuries caused by stroke or spinal cord lesions often have an impact on walking functions [37]. For example, stroke patients encounter sensory, motor, cognitive, and visual impairments which affect their ability to do many activities of daily life [9]. A majority of stroke patients experience motor impairments, e.g., limitation or loss of muscle coordination and strength. Motor deficits of the legs impact balance and walking ability. Hence, gait recovery is a crucial aspect of stroke rehabilitation [17]. Efficient rehabilitation requires repetitive and exhaustive exercises to induce neuro-plastic adaption and functional recovery [21]. However, the frequency and intensity of the traditional physical rehabilitation exercises are insufficient to achieve an effective recovery [9] due to a lack of eagerness and motivation [18].

Rehabilitation robots (see e.g., Fig. 2) increase the effectiveness of individual training sessions, hence their application has steadily increased [8]. Robot-assisted training usually implies high-dosage and high-intensity training [31]. Hence, robot-assisted therapy enhances the efficiency of individual training sessions. In addition, approaches using Virtual Reality (VR) have shown to moderately improve gait and balance rehabilitation, especially when combined with conventional rehabilitation methods (see [14, 21, 25] for recent meta-reviews). Still, the need for repetitive exercises remains, leading to the same problem of low motivation and frustration [37]. Here, VR approaches have shown to be capable of increasing motivation and decreasing the perception of exertion of patients [32]. VR is capable of simulating an almost endless variety of interesting artificial environments [13] during work-outs while it additionally provides

2019 IEEE Conference on Virtual Reality and 3D User Interfaces
23-27 March, Osaka, Japan
978-1-7281-1377-7/19/$31.00 ©2019 IEEE

Figure 2: Example of a gait robot used for clinical rehabilitation. Gait robots typically apply a cyclic movement that needs to be supported by the patients own gait force.

multi-modal sensory stimulation potentially beneficial for certain neuro-rehabilitation tasks. Notably, most prior work in this area used a somewhat restricted notion of VR (e.g., a low immersiveness), or did not control for long-time changes of important VR-based target effects of novelty and motivation.

Contribution: This paper reports on the design, development, and evaluation of an immersive VR-based rehabilitation system for neurological gait impairments applicable in a real-world clinical setup. The system provides a virtual environment consisting of open fields, forest, mountains, waterfalls, animals, and a beach for inspiring strolls and is able to include a virtual trainer as a companion during the walks.

We performed a three-day repetitive assessment and investigated the impacts of the VR medium as well as of the presence of a trainer on intrinsic motivation and the perceived user experience. We specifically investigated the effects using a longitudinal design, as single assessments might suffer from a casual temporal excitement, typically for VR experiences. We found notable effects. First, VR simulation could significantly increase enjoyment even though the enjoyment decreases throughout the study. Second, a physical trainer is more acceptable for individuals compared to the virtual trainer.

Structure: We continue with a review of the relevant related work in section 2. Section 3 then illustrates the design of the system. Section 4 describes the evaluation, design of the longitudinal study, the experimental procedures, and measures are taken. Section 5 presents the results which are followed by an interpretation and discussion in section 6. Finally, in section 7, a conclusion is drawn, and ideas and plans for future work are highlighted.

2 Related Work

Patients' motivation and engagement are essential factors in rehabilitation to improve performance and effectiveness [27]. Considerable research has been performed to investigate the effect of VR in therapy and users' acceptance.

Zimmerli et al. developed a VR application to foster motivational aspects in gait rehabilitation. The application consisted of different environments and tasks. The environments were displayed on a high-resolution 42-inch monitor which placed in front of the Lokomat. VR increased motivation and activity of the patients during training [37]. A study of Yang et al. illustrated that the integration of VR in the gait training program was useful and feasible for stroke patients and in addition, VR-based intervention is safe and encouraging. The virtual environment included different scenarios (e.g., lane walking, street crossing, obstacles striding across, and park stroll) and the stimulus displayed on three 239-cm wide connected screens in front of the subjects and an electronic system was used to track the legs movements and detect the collision with the virtual obstacle [36].

Bergmann et al. assessed the acceptability of robot-assisted gait training with/without VR and patients' motivation during the interventions. Two different scenarios (dog scenario and coin scenario) with various tasks were used, and the environment was, again, presented on a 42-inch screen in front of the patients. Results confirmed high acceptability of repetitive VR-augmented robot-assisted gait training and increased motivation for stroke patients [4]. Moreover, Richards and his colleagues reported improvement in mobility and higher self-efficacy in VR-coupled treadmill training. In this research, the subject walked with the preferred walking speed on a treadmill while looking at constant pictures.

Other approaches increased the immersiveness by projecting the VR environments on a big screen. In-line with our approach, Calogiuri et al. evaluated physical activity with and without a Head-Mounted Display (HMD). They used the Samsung gear 360° camera to record 360° video. This study reported cybersickness for the immersive VR condition and a higher enjoyment for the comparison condition of walking in the real environment [7].

Researchers at the University of North Carolina developed a rehabilitation framework including a virtual human physiotherapist. The virtual human indicates subjects how to do the bicep curl, and afterward, she asked the user to do the exercise for 30 seconds. During the training, the virtual human physiotherapist monitors the performance and gives feedback to the user. In addition, the virtual trainer shows the correct way of the exercise if the user performed the training incorrectly [1]. Study of Calabro et al. reported that VR and robotic-gait assisted device improved the gait and balance in patients with chronic hemiparesis. The VR application with agent projected on the 42-inch screen placed in front of the Lokomat with a 7.1 Dolby Surround system [6].

Shema and colleagues developed a VR Gait training system that incorporates treadmill training in a virtual environment. The VE consists of an obstacle course placed along different pathways. The VE was displayed on a screen in front of the treadmill. The results of 5 weeks clinical study revealed that the approach was effective and practical for patients with gait instability [30]. In a further study, Rooij et al. [29] performed the longitudinal pilot study with the post-stroke patients in 4 weeks. The virtual environment presented on a 180° semi-cylindrical screen. Participants performed exercises on the Gait Real-time Analysis Interactive Lab (GRAIL). The study reported that balance and gait ability improved after the patients experienced VR training. Furthermore, motivation and enjoyment increased.

The mirror neuron system has a significant role in human mimicking behavior and as such, relates to imitation and learning new proficiencies [28]. This part of the neural system activates when a person observes actions executed by another person [22]. Burns [5] reported that observing the motor acts by post-stroke patients may speed up their rehabilitation. Calabro and colleagues [6] discovered that integration of VR and robotic-based rehabilitation improved motor performance for the stroke patient. In fact, VR simulation influenced brain areas, especially mirror neurons, only caused by observing a walking activity.

2.1 Discussion

While IJsselsteijn et al. (2006) [15] could find a positive significant impact of higher immersion on performance, on intrinsic motivation and, on the presence, they did not validate and report the development of these effects over time. In contrast, Mestre et al. (2011) did check for effects over time [24]. They reported a significant decrease in exercise duration in a condition combining a stimulus of video and music. Here, a commitment was reduced with task repetition for a video-feedback-only condition but adding music listening resulted in a stable reported commitment across sessions. They conclude that neither the video nor the video and music condition represents a significant means to improve enjoyment during exercise performed with VR equipment. Notably, both approaches targeted indoor biking training for healthy users. Additionally, the VR-related qualities were limited: Field of view was the primary measure for immersion, feedback resulted in speed adoption of a virtual race track (which was either pre-recorded or generated during application). No stereoscopy or user-centered projection by head-tracking was employed.

Overall, most of the prior work used a restricted notion of VR (e.g., pre-recorded stimuli, low immersion, etc.), or did not control for long-time changes of important target effects or any combination of these restrictions. Each work motivates on a specific aspect and positively motivates the novel combination of the individual design and evaluation criteria of the work reported here.

3 Methods

We developed a virtual environment (VE) which consisted of four different nature landscapes: grassland, forest, stream land and beach (see Fig. 1). The individuals were immersed using a Head-Mounted Display (HMD) and asked to perform a typical stroke exercise on a cross-trainer. We further investigated the use of a virtual trainer as we expected that i) the trainer could provide additional motivation and ii) the trainer might provoke the activation of mirror neurons [22] if individuals observe the motion of the trainer and imitate a walking action.

3.1 Virtual Trainer

Visual presence and appearance of a virtual trainer can be important factors that impact motivation [2]. We designed the virtual trainer with a similar appearance to the experimenter. The female trainer walks in front of the user for the length of the simulation (see Fig. 3). In the beginning, the virtual trainer instructs the subject to walk 1000 steps and match the walking speed with the metronome sound. The virtual trainer increases the encouragement with motivational speeches (like "Nice Job!, Awesome! etc.") when the user walks 100, 250, 500, 750 and 1000 steps. In addition, the virtual trainer warns the subject when the speed of the movement is not corrected. The audio was based on recorded human voice, and the experimenter (physical trainer) recorded the motivational and warning speeches.

Figure 3: Virtual human trainer (left). A bubble speech in this figure is presenting the motivational speech, and it has not appeared in the VE. Virtual shoes were simulated to provide a simple form of embodiment to the user and to provide a reference point for their movements (right).

3.2 Apparatus

The simulation was developed using Unity engine 2017.4.0f1 running on a Windows 10. The character of the virtual trainer was created with Mixamo 3D animation software. The hardware setup consisted of one PC station (Intel Core i7-6700k 4.0 GHz CPU, 32 GB of RAM, NVIDIA GeForce GTX 1070 Graphics card). In all conditions, the participants walked using the Cardiostrong Cross trainer (EX90 PLUS). The participants visually immersed in a virtual environment using the HTC Vive stereoscopic Head-Mounted Display (HMD), with a field of view of 110 nominal, resolution of 2160 x 1200 pixels and refresh rate of 90Hz. Two wireless motion tracked controllers attached to the pedals of the cross trainer to measure data of the feet movements during walking activity (see Fig.4). The controllers were used to measure the number of steps, walking distance (m), average velocity (m/s), average steps per minute, maximum velocity (m/s) and maximum steps per minute. The axial ("in walking direction") distance between both controllers is used to calculate a step (i.e., the leading controller switches), and the maximum distance during that step is used as the step length. Step length and time of the last step are used to calculate, the velocity (m/s) and steps per minute. The step length in the study was physically fixed and the same for each participant. Participants used headphones which are attached to the auxiliary port of the HMD headset to hear the sound of metronome and environment in both VR conditions. In contrast, for the Non-VR conditions, the subjects heard the sound of the metronome from a single loudspeaker.

Figure 4: Handles of the cross trainer with the heart rate sensors (left). The subjects hold their hands between red tapes. Two wireless motion controllers are attached to the pedals (right).

4 Evaluation

4.1 Design

The longitudinal study conducted in a 2x2x3 mixed factorial design with the *medium* and the *presence of a trainer* as between factors, and the *day of assessment* as within factor. The four resulting conditions are depicted in Fig. 5. Participants were randomly assigned to one of the four conditions, and each subject repeated the same condition in three consecutive days.

4.2 Task

Participants walked 1000 steps with an average velocity of 0.27 m/s and average steps per minute of 39.69. Walking speed was adjusted based on metronome sound with a regular interval of 40 beats per minute (BMP). In the first day, each participant walked 30 seconds to adapt himself with the cross trainer. Moreover, each participant had another 30 seconds to acclimatize with the environment.

4.3 Measures

4.3.1 Control Measures

We assessed the participants' general motivation using the Global Motivation Scale (GMS) [10]. The GMS evaluates the general motivation in a trait measure using a 7-point scale from 1 (low) to 7 (high). The GMS assesses three types of intrinsic motivation (toward knowledge, stimulation, and accomplishment), three types of

Figure 5: Conditions of the experiment. From left to right: Non-VR no trainer, Non-VR with a trainer, VR no trainer, VR with a trainer.

extrinsic motivation (identified, introjected, and external regulation) and amotivation.

Intrinsic motivation to know (IM- to know) refers to an individual who performs a task for pleasure while exploring or learning new things. In the second types of intrinsic (IM-to accomplish), persons engaged in an activity for satisfaction when they try to accomplish something. An IM to stimulation sub-scale can be defined as engaging in an activity to encounter stimulating feelings.

Identified regulation happens when behaviour is valued, and individuals perform that behaviour because it is important for their personal growth. Introjected regulation is kind of motivation from pressuring voice and internalized. Individuals perform a task because of a sense of fear, shame or guilty [34]. External regulation consists of behaviours that perform to get a reward or prevent negative consequences [11].

Finally, amotivation occurs when the possibilities between actions and consequences do not realize by individuals [26]. There is a lack of motivation or intention to engage in a task and individuals cannot see the point in doing an activity [35].
Furthermore, we assessed how often participants perform sports along with the body mass index, and the usage of media along with previous experiences with VR. We further assessed quantitative simulation data, i.e., the average speed, pauses, and step frequency.

4.3.2 Subjective Measures

The Simulator Sickness Questionnaire (SSQ) [16] with a list of 16 symptoms was conducted before and after walking on the cross trainer to check the unwanted side-effects and cybersickness. Each symptom measured with the scale from non, slight, moderate to severe. The NASA Task Load Index (NASA-TLX) [12] was conducted to assess the perceived workload after walking activity on the cross trainer, the scale is divided by 21 vertical tick- marks into 20 equal intervals with ranges from 0 (low) to 100 (high). Intrinsic Motivation Inventory (IMI) [23] was used to evaluate the intrinsic motivation of the participants for walking activity. In other words, intrinsic motivation occurs when an individual engages in an activity for the satisfaction and enjoyment. In this study, we constructed the IMI questionnaire with 30 items out of 45 items with 5 scales namely: enjoyment, competence, effort, pressure, and value. The items scored on a 7-point Likert scale from 1(strongly disagree) to 7 (strongly agree). User effect was assessed with the International Positive and Negative Affect Schedule (PANAS) Short Form (I-PANAS-SF) [33], using a five-point Likert scale from 1 (low) to 5 (high). The User Experience Questionnaire (UEQ) [19] was conducted to assess the experience of the participants using a seven-point Likert scale from -3 (low) to 3 (high). The UEQ measures attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty.

4.3.3 Physiology

We assessed the heart rates of the participants after 0, 25, 500, and 1000 steps using the heart rate sensor at the handles of the cross trainer. We asked the participants to hold their hands between two red tapes during walking activity(see Fig. 4).

4.4 Procedure

Each subject participated in the study in three consecutive days, and for each day, he/she spent 75 minutes to complete the tasks (see Fig.6) In the first day of the study, an experimenter asked the participants to read instruction of the experiment and guidelines of using HTC Vive for those who assigned in VR conditions. For each day of the study, the participants read and signed a consent form before starting the experiment. The participants answered the demographic questions and GMS questionnaire only for the first day of the study. The experimenter instructed the participants to position on the cross trainer, placing their hands between two red tapes on the handle and feet on the pedals. In each condition, participants were instructed to walk on the cross trainer for 25 minutes (or 1000 steps). The participants synchronized their walking speed with beats of a metronome (40 beats per minute) which they heard during the walking exercise. Afterwards, the subjects completed the post-experimental questionnaires.

In the VR conditions, participants wore a HMD. Four virtual landscapes with similar arrangements were presented to the participants in three days. The transition between each landscape was seamless, i.e., nature changed from grassland to forest to stream land and to beach, see Fig. 1. The duration of exposure to grassland, forest, stream land and to the beach were 3, 5, 15, and 2 minutes respectively. The participants continuously immersed in the environment for 25 minutes and there was no break between each landscape. The participants walked through the whole the landscapes in each day. The experimenter informed the participants about having a break anytime, but none of the participants requested for a break. In both, the virtual and physical condition, the same motivational sentences and same instructions where used by the physical as well as the virtual trainer.

4.5 Participants

We recruited 45 student participants (22 females) that took part in the experiment on three consecutive days (allowing for day/weekend pause). The participants' age ranged from 19 to 26 years ($M = 21.82$, $SD = 1.84$). Tab. 1, presents distribution of the participants in each condition. In 9 cases, the participants asked the experimenters to change one day of the study, and as the study should perform in three consecutive days, we offered to them a day before or respectively after the day that they could not participate. No participant performed two trials a day. One of the participants had a knee surgery one year before the participation in the study, which did however not have an impact on the performance. All participants had normal to corrected-to-normal vision. Participants reported doing an average of $M = 3.26$ ($SD = 2.21$) hours of sports per week. A two-way analysis of variance (ANOVA) revealed that the reported sport exercise hours did not differ significantly between the conditions. The participants had a mean BMI of $M = 22.13$ ($SD = 2.45$). BMI did not significantly differ between the conditions. Further, None of the subscores of the global motivation scale differed between the conditions. Participants revealed an average of $M = 8.03$ hours of previous VR experience, which did not differ significantly across conditions. All participants in the VR conditions reported at least one previous VR experience.

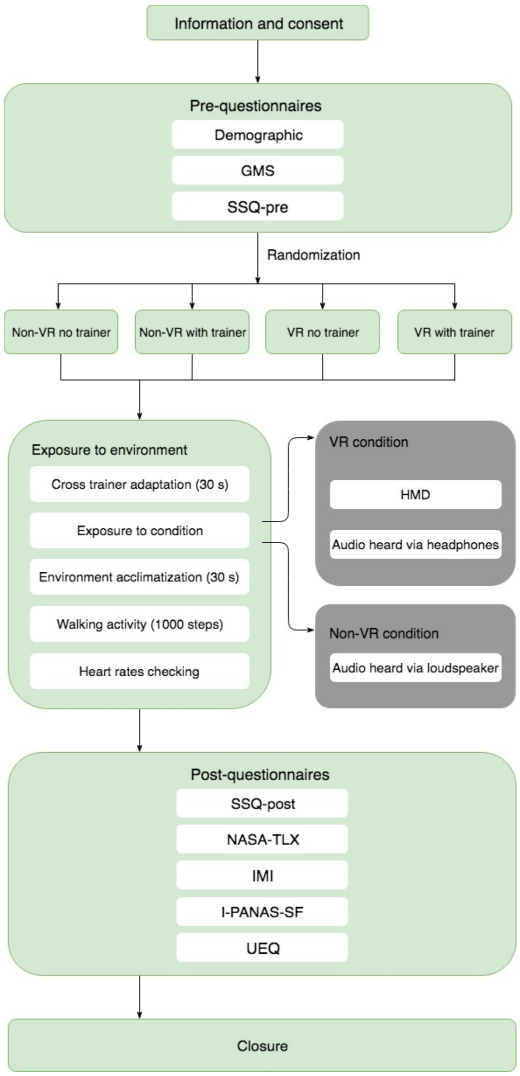

Figure 6: Experimental procedures. Each participant repeated these procedures for three days. The participant answered the demographic questions and GMS questionnaire on the first day of the study as well as adapting with the cross trainer.

5 RESULTS

A separate three-way mixed ANOVA was run to understand the effects of the *medium*, the *presence of a trainer* and the *day of assessment* for each dependent variables, using *medium* and the *presence of the trainer* as between factors, and the *day of assessment* as within factor, respectively. As the data was entered by the participants using a digital questionnaire, we could exclude data entry errors and measurement errors, and thus treated genuinely unusual values (outliers) conservatively by keeping them in the analysis. Partial η^2 (η_p^2) is reported as a measure of effect size. Where the assumption for sphericity was violated as assessed by Maulchy's test for sphericity, Greenhouse Geisser corrected values are reported. Pairwise comparisons report Bonferroni adjusted values.

5.1 Simulator Sickness

We analyzed Simulator Sickness using the method from Kennedy [16], calculating the differences of the total score between (post-pre).

We did not find any significant main or interaction effects. The difference score in the VR condition was slightly higher ($M = 6.32$, $SE = 2.65$) compared to the Non-VR condition ($M = 1.76$, $SE = 2.71$). Overall, the sickness difference was highest on day 3 ($M = 5.86$, $SE = 1.59$) compared to day 2 ($M = 3.37$, $SE = 1.68$) and day 1 ($M = 2.89$, $SE = 3.26$).

As we thought that due to the sports activity, the item for "sweating" might have biased these results, we excluded the item to substantiate the analyses. Again, there were no significant main or interaction effects. Difference scores were slightly higher in VR ($M = 3.58$, $SE = 2.54$) compared to the Non-VR condition ($M = 0.14$, $SE = 2.59$).

Table 1: Distribution of the participants across conditions

Condition	Female	Male	Total
Non-VR no trainer	6	5	11
Non-VR with trainer	5	6	11
VR no trainer	5	6	11
VR with trainer	6	6	12
Total	22	23	45

5.2 Task Load

To analyze the results, we assessed the RAW TLX scores for each measure.

5.2.1 Mental Demand

Regarding the mental demand participants required to perform the exercise, we found a significant main effect for the day of assessment $F(1.15, 47.11) = 9.54, p = .002, \eta_p^2 = .195$. Results indicate, that the mental demand was significantly higher on day 1 ($M = 14.68$, $SE = 2.27$), compared to day 2 ($M = 8.43$, $SE = 1.20$, $p = .022$) and day 3 ($M = 7.35$, $SE = 0.97$, $p = .002$). The difference between day 2 and day 3 was not significant.

5.2.2 Physical Demand

Regarding the physical demand participants required to perform the exercise, we found a significant main effect for the day of assessment $F(1.15, 60.54) = 7.60$, $p = .003, \eta_p^2 = .156$. Pairwise comparisons show that the physical demand was highest on day 1 ($M = 25.00$, $SE = 2.81$), compared to day 2 ($M = 21.34$, $SE = 2.08$) and day 3 ($M = 19.70$, $SE = 2.05$). The difference between day 2 and day 3 as well as between day 1 and day 3 was significant (p's $<= .026$). We further found a significant main effect for medium, $F(1, 41) = 5.01$, $p = .031$, $\eta_p^2 = .109$ indicating a higher perceived physical demand for the VR conditions ($M = 26.0$, $SE = 2.91$) compared to the Non-VR conditions ($M = 16.70$, $SE = 2.97$). Pairwise comparisons showed that this effect resulted mainly from the conditions where no trainer was present ($p = .047$) whereas the comparison of the trainer conditions was non-significant. Each day, both VR conditions were perceived higher in physical demand compared to the Non-VR conditions, but that this effect was only significant for day 1 ($p = .031$) where the highest difference was perceived ($M_d = 12.54$) whereas on day 2 ($M_d = 7.23$) and day 3 ($M_d = 8.12$) the difference was not significant.

5.3 Temporal Demand and Performance

We found significant main effects for the day of assessment for both temporal demand and performance, indicating that the temporal demand was lowest on day 1; $F(1.85, 75.80) = 5.11$, $p = .010, \eta_p^2 = .111$, as was the perceived performance; $F(1.24, 50.77) = 5.74$, $p = .015, \eta_p^2 = .123$.

5.4 Intrinsic Motivation

5.4.1 Enjoyment

We found a significant main effect for enjoyment for the day of assessment; $F(2, 82) = 5.05$, $p = .009$, $\eta_p^2 = .110$. As expected, enjoyment was highest on the first day ($M = 3.52$, $SE = 0.18$), compared to the second day ($M = 3.36$, $SE = 0.18$), and the third day ($M = 3.15$, $SE = 0.20$), see Fig.7. Enjoyment was higher in the VR conditions ($M = 3.63$, $SE = 0.25$) compared to the Non-VR conditions ($M = 3.06$, $SE = 0.25$). This general difference was not significant ($p = .121$). However, pairwise comparisons revealed that enjoyment was significantly higher for VR when no trainer was present, compared to the Non-VR condition ($p = .043$) see Fig. 7 .

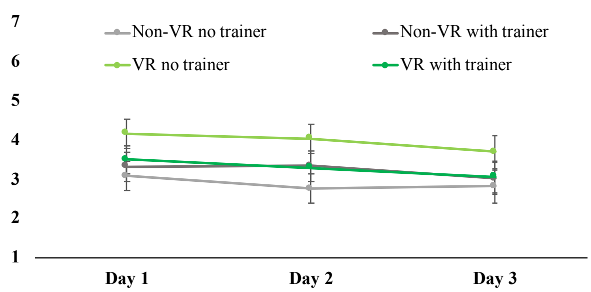

Figure 7: Mean enjoyment between the medium and day. The enjoyment decreased throughout the study.

5.4.2 Competence and Effort

There was a main effect for both, competence and effort for the day of assessment; $F(2, 82) = 11.88, p < .001, \eta_p^2 = .225$. Competence was higher on the second day ($M = 5.42$, $SE = 0.13$) compared to the first day ($M = 4.94$, $SE = 0.14$; $p < .001$) but slightly decreased on the third day ($M = 5.36$, $SE = 0.14$). The perceived effort was highest on day 1 ($M = 4.82$, $SE = 0.14$) and decreased in day 2 ($M = 4.77$, $SE = 0.14$) and day 3 ($M = 4.35$, $SE = 0.15$) compared to day 1 (p's $<= .001$); $F(1.68, 68.65) = 8.76$, $p = .001$, $\eta_p^2 = .176$.

5.4.3 Pressure

The perceived pressure was significantly higher in both VR conditions ($M = 2.33$, $SE = 0.14$) compared to the Non-VR conditions ($M = 1.78$, $SE = 0.14$; $F(1, 41) = 7.65$, $p = .008$, $\eta_p^2 = .157$). Pressure was highest on day 1 ($M = 2.22$, $SE = 0.13$) and lowest on day 3 ($M = 1.91$, $SE = 0.93$); $F(2, 82) = 4.39$, $p = .015, \eta_p^2 = .097$.

5.4.4 Value

We found a significant interaction effect for *medium* X *presence of a trainer* for the perceived value; $F(1, 41) = 10.05$, $p = .003, \eta_p^2 = .197$. Comparisons show that the VR condition was rated higher when no virtual trainer was present ($M = 4.25$, $SE = 0.32$) and lower when a virtual trainer was present ($M = 3.53$, $SE = 0.31$). Vice versa, the Non-VR condition without physical trainer ($M = 3.25$, $SE = 0.32$) was rated lower compared to the Non-VR condition with physical trainer ($M = 4.53$, $SE = 0.32$), see Fig. 8. Therefore, the physical trainer increased the value of the exercise whereas the virtual trainer decreased the perceived value.

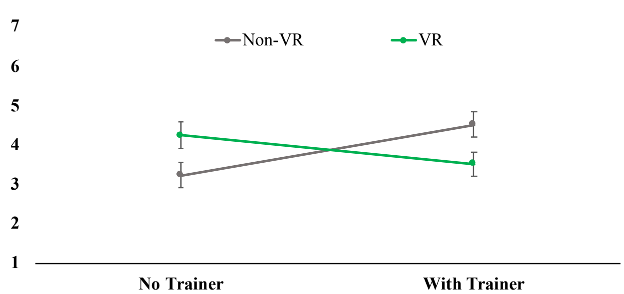

Figure 8: Interaction between the medium and the presence of a virtual trainer. While in the non-VR condition, the physical trainer increased the value of the exercise, the virtual trainer in the VR condition decreased the perceived value.

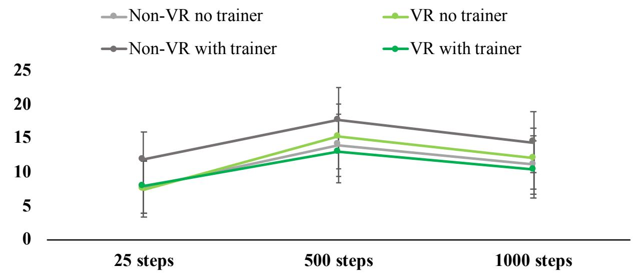

Figure 9: Mean heart rate differences from the baseline throughout the step measurements at respectively 25, 500, 1000 steps.

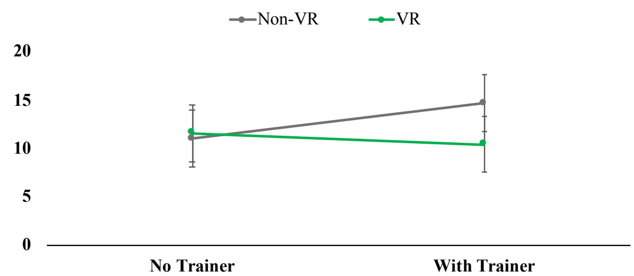

Figure 10: Mean heart rate differences of the 4 conditions from the baseline averaged over all time points.

5.5 Affect

There was a significant main effect for affect in for the day of assessment; $F(2, 82) = .10.09$, $p < .001$, $\eta_p^2 = .197$. Positive affect decreased significantly from day 1 ($M = 3.14$, $SE = .114$) to day 2 ($M = 2.86$, $SE = .089$) as well as from day 1 to day 3 ($M = 2.71$, $SE = .110$, $p < .001$).

5.6 Heart Rate

To measure the physiological response of the exercise, we calculated difference scores from the measures at 25 steps, 500 steps, and 1000 steps to the baseline at 0 steps. This time point was added as additional within-factor for the ANOVA. We found a significant main effect for time point (step number) of the assessment; $F(2, 82) = 15.38$, $p < .001$, $\eta_p^2 = .273$. The heart rate increased between 25 and 500 steps and decreased from 500 to 1000 steps. Figure 9 shows the absolute difference measures from the baselines. While in VR condition, the trainer had a slight decreasing effect on the heart rate, the trainer had an increasing effect on the heart rate in the Non-VR condition, see Fig. 10.

5.7 User Experience

5.7.1 Perspicuity

There was a significant main effect for day regarding perspicuity; $F(2, 41) = 3.30$, $p = .042, \eta_p^2 = .074$; Perspicuity increased from day one ($M = 6.17$, $SE = 0.11$) to day 2 ($M = 6.37$, $SE = 0.77$) and was then stable ($M = 6.36$, $SE = 0.89$). There was a significant main effect for medium regarding perspicuity; $F(1, 41) = 4.90$, $p = .033$, $\eta_p^2 = .106$; The VR conditions were rated significantly lower regarding perspicuity ($M = 6.13$, $SE = 0.11$) compared to the Non-VR conditions ($M = 6.47$, $SE = .112$).

5.7.2 Novelty

There was a significant main effect for medium regarding novelty; $F(1, 41)= 69.41$, $p < .001$,$\eta_p^2 = .629$. VR was rated significantly higher in novelty ($M = 5.06$, $SE = 0.21$) compared to the Non-VR condition ($M = 2.52$, $SE = 0.22$). Pairwise comparisons revealed that whereas the Non-VR conditions had a low but stable novelty rating, the VR conditions significantly decreased in novelty from day 1 ($M = 5.31$, $SE = 0.224$) to day 2 ($M = 4.97$,$SE = 0.24$;$p = .031$) and was stable for day 3 ($M = 4.89$, $SE = 0.23$).

5.7.3 Stimulation

There was a significant interaction for *medium* X *presence of a trainer* for the stimulation measure; $F(1, 41)= 4.52$, $p = .039$,$\eta_p^2 = .099$. Pairwise comparisons showed that without the presence of a trainer, the VR condition ($M = 4.35$, $SE = 0.32$) was rated significantly higher than the Non-VR condition ($M = 3.13$, $SE = 0.32$), whereas with the presence of a trainer, the comparison between VR ($M = 3.92$, $SE = 0.31$) and the Non-VR condition ($M = 4.05$, $SE = 0.32$) was non significant, even showing higher values for the Non-VR condition. In the Non-VR condition, the trainer increased stimulation significantly ($p = .049$). We further found a significant main effect for day $F(1, 41)= 5.03$, $p = .009$, $\eta_p^2 = .109$. Stimulation decreased for all consecutive days in both for both media. Without the presence of a trainer, stimulation also decreased throughout each day, but the presence of a trainer could stop this trend for day three ($M = 3.85$,$SE = 0.24$) compared to day two ($M = 3.81$, $SE = 0.24$). Exploratory comparisons for each day further revealed that the stimulation of VR was only significantly higher for the first day ($p = .046$) whereas the comparisons of the consecutive days were non-significant.

5.7.4 Efficiency

The interaction of *medium* X *presence of a trainer* was significant for the efficiency rating; $F(1,41) = 7.66$, $p = .008$, $\eta_p^2 = .157$. Pairwise comparisons showed that the Non-VR condition was rated significantly more efficient with the presence of a trainer ($M = 5.02$, $SE = 0.17$) compared to the VR condition with the presence of a trainer ($M = 4.25$, $SE = 0.16$;$p = .002$). The comparison between the media without a trainer was non significant. We further found an interaction between the the *day of measurement* and the *presence of a trainer*; $F(2,82) = 7.49$, $p = .001$,$\eta_p^2 = .154$.

5.7.5 Dependability

There was a significant main effect for the day of measure for the perceived dependability $F(2,82) = 7.82$, $p = .001$, $\eta_p^2 = 0.160$. Dependability was lowest for day 1 ($M = 5.70$, $SE = 0.12$) and highest for day 3 ($M = 6.04$, $SE = 0.84$). We further found a significant main effect for *medium*; $F(1,41) = 12.87$, $p = .001$, $\eta_p^2 = 2.39$. Both Non-VR conditions ($M = 6.22$, $SE = 0.12$) were rated significantly higher in dependability compared to the VR conditions ($M = 5.61$, $SE = 0.12$, $ps < .024$).

No further significant main or interaction effects were found in the analyses.

6 Discussion

As expected, the enjoyment of the exercise decreased throughout the day of assessment (see Fig.7). While there was a significant increase in enjoyment for VR, this effect was not a general effect but limited to the conditions without a virtual trainer. We think that a rather simple implementation of the virtual trainer is not able to compete with the socio-motivational impacts of a physical trainer and thus did hinder an increase in enjoyment in the VR condition. In addition, the participants could see the face of the virtual trainer from the side when she said the motivational sentences. Further research should improve the capabilities of the virtual agent to increase interactivity and humanness. We figured out that effort was mainly higher in day 1 compared to day 2 and day 3. This could be explained that on the first day the participants performed more tasks and walking activity compared to the second and third day. In addition, none of the participants used their breaking time after walking activity. Our finding suggests that, for further research, we need to ask the participants to rest during the study.

As one important aspect, we found a significant interaction in the perceived value, showing that the virtual trainer decreased the exercise value whereas the physical trainer increased the exercise value. To that regard, it is important to note that while we tried to control both conditions as good as possible, our virtual trainer did not have any special social responses or richness of social cues or behaviors. The actions of the trainer were limited to a simple head tilt (about 30deg) accompanying the motivational sentences. This finding can, therefore, be interpreted in a way that participants rated based on a social value they assign to a human companion, compared to a virtual companion. While this rating is explainable and rational, future research should improve the agent's social skills and behaviors. For example, researchers found that for a motivational agent facial expressions (eye and mouth movements) and deictic gestures (e.g., pointing with hands) are beneficial, in order to deliver the motivational message to the subjects [3]. However, we think that this finding underlines the fact that while virtual rehabilitation can be an addition to traditional rehabilitation, the human factor is still important for the users. This is also reflected upon by the dependability scores.

The outcome revealed that VR conditions were significantly higher in novelty. However, the novelty perception, most rationally, dropped after day 1. This could be explained due to (1) implementing the longitudinal assessment. (2) The individuals were exposed to the same VR stimuli with a similar task. We may argue that VR can be attractive for the individuals at the beginning but not for a long time especially if the individuals were exposed to the same VR stimuli. We should consider changing of VE for future work. In addition, conforming to Yan [36] and Shema [30], we should provide different tasks with varying levels of complexity. We found significantly higher stimulation and efficiency of Non-VR condition with the trainer.

The results show a considerable effect of the heart rate for a different number of steps. The heart rate continuously increased until 500 steps for all the conditions and then decreased. Interestingly, the heart rate declined in VR with a virtual trainer, whereas it increased in the non-VR condition with a physical trainer. This could either imply that the VR-based rehabilitation system had a relaxing impact on the participants [20]. However, taking into account the perceived value, another interpretation is that the human-to-human social connection that was present with a physical trainer (see also Fig 9) led to a higher physiological activation. Further research is necessary to gain more insights.

In contrast to the study of Calogiuri [7] that assessed VR with healthy participants in one day, we found that simulator sickness was not increased significantly due to the VR simulation in this longitudinal study. Therefore, despite the fact that the participants had a physical coupling to the cross trainer, there were no significant sickness effects. We believe that the main reason for this is the continuous slow movement of the patients in the virtual world with a focus in the forward direction. As the exposure to this stimuli was rather high, we assume an acclimatization process. We thus evaluate the simulation applicable for further clinical trials. Interestingly, we did find a lower physical demand in the physical condition compared to the virtual condition, especially without the presence of a trainer. The first interpretation of this finding might be that the users perceived the virtual simulation, i.e., the walk through the environments and the distance performed as a long distance, whereas a relation in the physical world was not given. The second interpretation of higher demand in VR might be the wearing of additional sensors. The trainer conditions did not show a high difference. Additionally, we found that the position of the physical trainer was next to the participant, but the position of the virtual trainer was in front of the participants. This point should be considered for further study.

7 Conclusion and Future Work

We developed the VR-based rehabilitation system which provided a virtual environment with different landscapes and a virtual trainer. We performed the longitudinal study with healthy participants to evaluate the system. We assessed enjoyment, motivation and user experience. While we found a multitude of effects that developed over time, such as decreasing enjoyment, both the medium and the presence of the trainer had impacted the results. Participants attributed less value to a virtual trainer compared to a physical trainer. On the other hand, enjoyment was rated higher in the VR conditions compared to non-VR conditions. Importantly, we did not find an unwanted side effect for this longitudinal study, and thus we judge our simulation feasible for clinical trials. The results of this study helped us to figure out the requirements that are required to improve our system, such as to enhance the virtual trainer in terms of social communication, interactivity, and humankind.

References

[1] S. Babu, C. Zanbaka, J. Jackson, T. Chung, B. Lok, M. C. Shin, and L. F. Hodges. Virtual human physiotherapist framework for personalized training and rehabilitation. In *Graphics Interface*, pp. 9–11, 2005.

[2] A. L. Baylor. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3559–3565, 2009.

[3] A. L. Baylor. The design of motivational agents and avatars. *Educational Technology Research and Development*, 59(2):291–300, 2011.

[4] J. Bergmann, C. Krewer, P. Bauer, A. Koenig, R. Riener, and F. Müller. Virtual reality to augment robot-assisted gait training in non-ambulatory patients with a subacute stroke: a pilot randomized controlled trial. *European journal of physical and rehabilitation medicine*, 2017.

[5] M. S. Burns. Application of neuroscience to technology in stroke rehabilitation. *Topics in stroke rehabilitation*, 15(6):570–579, 2008.

[6] R. S. Calabrò, A. Naro, M. Russo, A. Leo, R. De Luca, T. Balletta, A. Buda, G. La Rosa, A. Bramanti, and P. Bramanti. The role of virtual reality in improving motor performance as revealed by eeg: a randomized clinical trial. *Journal of neuroengineering and rehabilitation*, 14(1):53, 2017.

[7] G. Calogiuri, S. Litleskare, K. A. Fagerheim, T. L. Rydgren, E. Brambilla, and M. Thurston. Experiencing nature through immersive virtual environments: Environmental perceptions, physical engagement, and affective responses during a simulated nature walk. *Frontiers in psychology*, 8:2321, 2018.

[8] W. H. Chang and Y.-H. Kim. Robot-assisted therapy in stroke rehabilitation. *Journal of stroke*, 15(3):174, 2013.

[9] I. J. de Rooij, I. G. van de Port, and J.-W. G. Meijer. Effect of virtual reality training on balance and gait ability in patients with stroke: systematic review and meta-analysis. *Physical therapy*, 96(12):1905–1918, 2016.

[10] F. Guay, G. A. Mageau, and R. J. Vallerand. On the hierarchical structure of self-determined motivation: A test of top-down, bottom-up, reciprocal, and horizontal effects. *Personality and social psychology bulletin*, 29(8):992–1004, 2003.

[11] F. Guay, R. J. Vallerand, and C. Blanchard. On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (sims). *Motivation and emotion*, 24(3):175–213, 2000.

[12] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[13] M. K. Holden. Virtual environments for motor rehabilitation. *Cyberpsychology & behavior*, 8(3):187–211, 2005.

[14] M. C. Howard. A meta-analysis and systematic literature review of virtual reality rehabilitation programs. *Computers in Human Behavior*, 70:317–327, 2017.

[15] W. A. IJsselsteijn, Y. d. Kort, J. Westerink, M. d. Jager, and R. Bonants. Virtual fitness: stimulating exercise behavior through media technology. *Presence: Teleoperators and Virtual Environments*, 15(6):688–698, 2006.

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

[17] P. Langhorne, F. Coupar, and A. Pollock. Motor recovery after stroke: a systematic review. *The Lancet Neurology*, 8(8):741–754, 2009.

[18] M. A. Latif, H. M. Yusof, S. Sidek, M. Shikhraji, and M. Safie. A gaming-based system for stroke patients physical rehabilitation. In *Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on*, pp. 690–695. IEEE, 2014.

[19] B. Laugwitz, T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*, pp. 63–76. Springer, 2008.

[20] K. Laumann, T. Gärling, and K. M. Stormark. Selective attention and heart rate responses to natural and urban environments. *Journal of environmental psychology*, 23(2):125–134, 2003.

[21] K. R. Lohse, C. G. Hilderman, K. L. Cheung, S. Tatla, and H. M. Van der Loos. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PloS one*, 9(3):e93318, 2014.

[22] F. Maeda, G. Kleiner-Fisman, and A. Pascual-Leone. Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation. *Journal of neurophysiology*, 87(3):1329–1335, 2002.

[23] E. McAuley, T. Duncan, and V. V. Tammen. Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport*, 60(1):48–58, 1989.

[24] D. R. Mestre, C. Maïano, V. Dagonneau, and C.-S. Mercier. Does virtual reality enhance exercise performance, enjoyment, and dissociation? an exploratory study on a stationary bike apparatus. *Presence: Teleoperators and Virtual Environments*, 20(1):1–14, 2011.

[25] D. C. Porras, P. Siemonsma, R. Inzelberg, G. Zeilig, and M. Plotnik. Advantages of virtual reality in the rehabilitation of balance and gait: Systematic review. *Neurology*, 90(22):1017–1025, 2018.

[26] C. F. Ratelle, F. Guay, R. J. Vallerand, S. Larose, and C. Senécal. Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of educational psychology*, 99(4):734, 2007.

[27] A. S. Rizzo and G. J. Kim. A swot analysis of the field of virtual reality rehabilitation and therapy. *Presence: Teleoperators & Virtual Environments*, 14(2):119–146, 2005.

[28] G. Rizzolatti and L. Craighero. The mirror-neuron system. *Annu. Rev. Neurosci.*, 27:169–192, 2004.

[29] I. Rooij, I. van de Port, and J. Meijer. Feasibility and effectiveness of virtual reality training on balance and gait recovery early after stroke: A pilot study. *Int J Phys Med Rehabil*, 5(417):2, 2017.

[30] S. R. Shema, M. Brozgol, M. Dorfman, I. Maidan, L. Sharaby-Yeshayahu, H. Malik-Kozuch, O. Wachsler Yannai, N. Giladi, J. M. Hausdorff, and A. Mirelman. Clinical experience using a 5-week treadmill training program with virtual reality to enhance gait in an ambulatory physical therapy service. *Physical therapy*, 94(9):1319–1326, 2014.

[31] M. Sivan, R. J. O'Connor, S. Makower, M. Levesley, and B. Bhakta. Systematic review of outcome measures used in the evaluation of robot-assisted upper limb exercise in stroke. *Journal of Rehabilitation Medicine*, 43(3):181–189, 2011.

[32] H. Sveistrup. Motor rehabilitation using virtual reality. *Journal of neuroengineering and rehabilitation*, 1(1):10, 2004.

[33] E. R. Thompson. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (panas). *Journal of cross-cultural psychology*, 38(2):227–242, 2007.

[34] R. J. Vallerand, L. G. Pelletier, M. R. Blais, N. M. Briere, C. Senecal,

and E. F. Vallieres. The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and psychological measurement*, 52(4):1003–1017, 1992.

[35] S. P. Vlachopoulos and C. I. Karageorghis. Interaction of external, introjected, and identified regulation with intrinsic motivation in exercise: relationships with exercise enjoyment. *Journal of Applied Biobehavioral Research*, 10(2):113–132, 2005.

[36] Y.-R. Yang, M.-P. Tsai, T.-Y. Chuang, W.-H. Sung, and R.-Y. Wang. Virtual reality-based training improves community ambulation in individuals with stroke: a randomized controlled trial. *Gait & posture*, 28(2):201–206, 2008.

[37] L. Zimmerli, A. Duschau-Wicke, A. Mayr, R. Riener, and L. Lunenburger. Virtual reality and gait rehabilitation augmented feedback for the lokomat. In *Virtual Rehabilitation International Conference, 2009*, pp. 150–153. IEEE, 2009.