Exergaming: The Impact of Virtual Reality on Cognitive Performance and Player Experience

1st Felix Born
Entertainment Computing Group
University of Duisburg-Essen
Duisburg, Germany
felix.born@uni-due.de

2nd Linda Graf
Entertainment Computing Group
University of Duisburg-Essen
Duisburg, Germany
linda.graf@uni-due.de

3rd Maic Masuch
Entertainment Computing Group
University of Duisburg-Essen
Duisburg, Germany
maic.masuch@uni-due.de

***Abstract*—Exergames can increase cognitive performance compared to sedentary games and the same non-gamified exercise. Whether this advantage can be further enhanced by the high immersion and arousal level VR systems offer remains unclear. Therefore, we developed an exergame that can be played in VR or in front of a TV screen with identical gameplay in both versions. In a study with 32 participants, we collected data on cognitive performance, arousal, presence, motivation, and exertion. Our results show no difference in cognitive performance and arousal between the conditions. However, the VR version significantly increased players' presence and motivation while significantly reducing perceived exertion. Our results help to understand in which aspects VR can have a significant impact on exergames.**

***Index Terms*—Virtual Reality, VR, Exergaming, Exergames, Exertion Games, Motivation, Cognitive Performance, Arousal**

I. Introduction

Exertion games or *exergames* are digital games "*[...] where the outcome of the game is predominantly determined by physical effort*" [31]. The use of exergames has demonstrated a positive impact in a variety of different domains, such as prevention and counteraction of childhood obesity [18], supporting physical rehabilitation [4] and improving cognitive performance [17].

While most of the exergaming research used older TV screen systems such as Microsoft's Kinect or Nintendo's Wii, current work focuses specifically on modern VR systems such as the HTC Vive. Since body movements serve as the main interactions, modern VR systems have an inherent exergaming property. Furthermore, first studies observed that in contrast to conventional non-VR exergames, VR exergames can improve central player experience related constructs such as motivation [6], [28] due to the high level of immersion VR systems offer. Thus, VR exergames are used for different purposes such as improving physical performance [2], decreasing perceived exertion [43], or supporting physical rehabilitation [19].

Beyond advantages that are directly linked to the context of physical activity, recent works used VR exergames to improve the cognitive performance of players [21], [26], [35]. For non-VR exergames, it has already been shown that exergaming can lead to better cognitive performance than playing a sedentary game [17] and performing the same gamified exercise [1]. Beyond that, the use of VR exergames can lead to enhanced cognitive performance compared to the same non-exergame physical exercise without VR [26], [35]. However, if the use of VR can enhance the cognitive benefits non-VR exergames provide remains unclear.

Therefore, we developed an exergame that uses either a conventional TV screen or an HMD as the output device but offers the same gameplay regardless of the output device. In a between-subjects design study with 32 participants, we investigated whether the VR exergame version could enhance the players' cognitive performance after a one-time training. In addition to cognitive performance, we measured the players' perceived and actual exertion, arousal, presence, and motivation to investigate their moderating effect on cognitive performance. Further, we used these constructs to investigate the impact of VR on player experience to support the research area with results from a between-subjects design study.

We did not observe a significant difference in arousal or cognitive performance between the conditions. Hence, we assume that the cognitive benefits exergames offer after a one-time training cannot be increased using VR. However, our results indicate that the VR version could significantly increase the players' motivation and presence while significantly reducing perceived exertion, although, as intended, actual exertion did not differ between the conditions. With our results, we address game developers and researchers alike. Our results support the insights on the comparison between VR and non-VR exergames. VR exergames but also conventional exergames in front of the TV (e.g., Nintendo's Ring Fit Adventure) are constantly developed further. Our results help identify areas in which VR systems can offer substantial benefits but also areas where VR may not further enhance the benefits exergames provide.

II. Immersive VR Exergames

Immersion can be defined as the extent of a user's technical integration into a virtual environment and is described as an objective characteristic of a technology [37]. Thus, modern

978-1-6654-3886-5/21/$31.00 ©2021 IEEE

room-scale VR systems offer a higher level of immersion compared to conventional screen systems. Immersion has a direct impact on the user's experience of presence [38] which can be defined as the feeling of "*being there*" [20] or simply as *presence*. In the context of exergames, a high sense of presence can significantly improve the player experience in various aspects.

VR exergames can significantly increase the player's experience of presence, self-efficacy, performance, and intrinsic motivation [6], [28] compared to non-VR exergames and to the non gamified version of the exercise [28], [43]. The player's intrinsic motivation, in particular, is a central construct in the context of VR exergaming since being intrinsically motivated can result in engaging in and enjoying a specific activity [34]. Moreover, on the one hand, VR exergames can reduce perceived exertion when used to distract from strenuous and repetitive exercises such as rowing or cycling [28], [43]. On the other hand, VR exergames that are driven by more complex and constantly changing interactions and thus do not require an exercycle or rowing machine can produce high motivation scores despite unreduced perceived exertion scores [6], [7], [16].

A. Exergaming, Arousal, and Cognitive Performance

However, beyond motivational benefits, which are directly related to the athletic context and the main focus of most studies on VR exergaming, exergames offer further benefits. One of these benefits is the improvement of cognitive performance through exergames [1], [17]. The positive influence of physical activity on cognitive performance has already been shown in various studies for a large number of different activities, intensities, periods, and target groups [11]. However, cognitive performance does not depend directly on physical activity but is influenced by arousal, which, among other things, is influenced by physical activity [24]. Arousal is defined as one of three dimensions of emotions besides valence [33] and dominance [29] and describes the excitation level of an affective state that ranges from calm to excited. In addition to physical activity influencing the arousal level, arousal can also be influenced by affective stimuli, such as pictures [25], sounds [9] or even game elements [30].

Beyond simple affective stimuli, the authors of [17] compared an exergame to the same game played sedentarily and to a conventional treadmill exercise. Although the exergame and the treadmill exercise induced similar physical exertion, the exergame induced more arousal. However, though the exergame showed improved cognitive performance compared to the sedentary game, no difference in cognitive performance to the treadmill condition could be found. Beyond that, the authors of [1] compared a three-month exergame intervention with a non-gamified version of the same athletic exercise. Although physical exertion did not differ between groups, the exergaming intervention significantly increased participants' cognitive performance. Unfortunately, the authors did not measure the participant's arousal as a potential moderator for cognitive performance. However, the results of [1] suggest that if the arousal level increases while physical exertion remains constant, the cognitive performance can be improved even more. With regards to the results of [17] and [1], the question arises how the arousal of exergames can be increased even further and whether an increase in arousal can benefit the cognitive performance even more, since for complex tasks, increasing arousal levels can decrease the cognitive performance after a certain arousal level is exceeded [42].

One way to further increase the arousal level in exergames is the use of VR. Several studies [10], [15], [32] observed increased arousal levels in VR games compared to non-VR games. Therefore, we hypothesize that the combination of VR and exergames can increase arousal through its physical exertion and high level of immersion. Thus, VR exergames have a potentially more substantial influence on cognitive performance compared to non-VR exergames. The influence of VR exergames on cognitive performance has been investigated in very few studies so far. The studies devoted to this topic have compared the effect of a VR exergame intervention with a control group that did not consist of a gamified version of the exercise. The authors of [35] compared a conventional training on a cycle ergometer to an exergaming version where the participants wore a Samsung Gear VR. Similar to the results of [1], the exergaming variant could lead to improved cognitive performance. However, the authors only compared a VR exergaming variant to the conventional non-gamified exercise but not a non-VR exergaming variant. Furthermore, the authors only measured cognitive performance and not arousal. In another work, the authors of [26] observed a positive effect of a VR exergame on older adults' cognitive performance above the age of 65 after a four-week intervention in contrast to no intervention.

The comparison between a VR and a non-VR version of an exergame regarding the effects on cognitive performance has been investigated by only one work so far. A recent study [21] showed that a VR exergame could lead to greater cognitive performance in people over 50 than a non-VR version of the game, but only after regular training over four weeks. However, the authors did not record the physical exertion as well as the arousal of the participants. Thus, it is difficult to determine whether the improved cognitive performance is due to the VR system's property or whether, for example, the physical exertion in the VR variant was higher than in the non-VR variant. Since the mere use of VR in the context of exergames can lead to increased exertion [41], it could be that the use of VR in the work of [21] increased physical exertion and thus influenced the cognitive performance. Furthermore, since only persons over 50 years of age participated in the study, it is particularly interesting to conduct further evaluations with a different age group,

Fig. 1. The sword is 71 cm long, 16 cm wide, and weighs 356 grams with the HTC Vive Tracker attached. The shield has a diameter of 50 cm and weighs 829 grams, including the mount and the HTC Vive Tracker. While the sword is held in hand, there is a mount for the entire forearm on the shield's back.

taking into account physical exertion and arousal. Moreover, since the author of [21] could not observe any effects for one-time training on the cognitive performance, but the work of [17] and of [35] could, the question arises whether the non-observable difference in the work of [21] is due to the high average age of the participants or the difference in the conditions compared.

B. Contribution

To gain a detailed insight into whether VR in the context of exergames can increase cognitive performance after a single training session and which factors can influence this, we developed an exergame that differs only in the output medium (VR vs. TV). To systematically evaluate which constructs are moderating this process, we record the constructs presence, motivation, exertion, and arousal in addition to cognitive performance in a study with 32 subjects. On the one hand, our results help to clarify the role of VR in enhancing cognitive performance in the context of exergames. Beyond that, our results also support the still relatively sparse research field comparing VR vs. non-VR exergames concerning central motivational and player experience-related factors.

III. VR Testbed Game

To investigate the difference between a VR and a non-VR exergame regarding their impact on the player's cognitive performance, we developed a new exergame with the game engine Unity. The game's setting takes the player to a small medieval village that Vikings have just plundered. As shown in figure 2, the player is standing at the end of a footbridge that leads from the village to a boat located behind the player. Vikings run toward the player to get on the boat and secure their loot. Besides, the opponents have occupied several positions in the village to throw skulls at the player. The players have to defeat the enemies running towards them and repel the skulls. For this, the players are equipped with a shield and a sword.

As shown in figure 1, we used padded weapons from the live-action role-playing domain for the shield and the sword. We mounted an HTC Vive tracker to both objects, using a clamp for the shield and a unique 3D-printed mount for the sword. The weapons are the only interaction objects in the game. The virtual representation of the weapons corresponds to the real ones. Since we wanted to develop a game that could be played in VR and in front of a conventional screen, the range of movements is severely limited if we want to compare the two conditions. However, to offer players a certain amount of exertion and interesting gameplay despite this limitation, we decided to use the real-looking interaction objects instead of lighter conventional controllers. Furthermore, in the context of VR, exergames using heavy controllers can increase the player's exertion while maintaining a high intrinsic motivation [7]. Since the interaction objects are mostly made of foam, they can be played with safely since even accidental hits of objects and body parts do not cause significant damage.

To keep the necessary movements identical in the VR and non-VR version, the players must perform all interactions with the shield and sword on the spot and do not have to walk across the room or rotate their vision. Hit enemies increase the player's score, while missed enemies decrease the score, which is displayed at a central location in the player's field of view, as shown in figure 2. To provide another motivational incentive, the player's striking power and the trajectory of defeated opponents continuously increase as the player consecutively defeats opponents. This mechanism is based on the idea of virtual performance augmentation from [22] and can increase intrinsic motivation in the context of VR exergames. The time intervals in which the enemies spawn shorten during the game time. Thus, the necessary exertion increases continuously. We have developed two different versions of the exergame, which only differ in the output device. In the VR condition, an HTC Vive Pro is used, while in the TV condition, a conventional Full HD 50 inch TV screen is used. We used a standard computer with a Ryzen 7 2700 CPU, 16 GB Ram, and an RTX 2070 Super GPU for both versions of the game. The gameplay and the interaction objects are identical in both conditions. Thus, we also use the HTC Vive trackers in the TV condition.

IV. Study

We conducted a quantitative study to evaluate the difference between the two versions of our exergame on cognitive performance and player experience.

A. Method

Cognitive performance was measured using the Stroop task [27]. In that task, participants are presented color words

Fig. 2. A screenshot of the gameplay from the player's perspective.

printed either in color the term stands for (congruent stimuli) or in a color that does not match the color's name (incongruent stimuli). Participants have to name the color the word is printed in and not the word itself. We used a computerized Stroop test provided by the *psytoolkit.org* website [39], [40]. Here, 40 color words were presented successively, and the participants had to press the corresponding button on the keyboard (e.g., for blue, they press "b"). We recorded the total reaction time, the reaction time for congruent and incongruent stimuli in seconds, and the number of errors. The arousal was measured by the Self-Assessment Manikin (SAM) [8] which includes one item for each affective state dimension (arousal, valence, dominance) that could be rated from 1 - 9.

To determine the perceived exertion, we used the Borg scale [5]. The scale asks for the sensation of exertion, which refers to the strain and fatigue of the muscles and the feeling of being out of breath and ranges from 6-20 (none - very, very hard). The actual exertion was assessed using the subjects' average heart rate during the baseline and gameplay phase. We used the *Bittium Biosignals Ltd. Faros 180* pulse belt. We used the Intrinsic Motivation Inventory (IMI) to measure the player's intrinsic motivation. The entire inventory comprises six subscales, of which the *Interest/Enjoyment* is considered the most significant one since it directly measures intrinsic motivation [14]. Beyond the *Interest/Enjoyment* subscale, we also assessed the *Perceived Competence* subscale and the *Pressure/Tension* subscale. Each subscale ranges from 1 - 7. The experience of presence was measured using the *IGroup Presence Questionnaire* (IPQ) [36]. The IPQ comprises the three subscales *Spatial Presence, Experienced Realism, Involvement* and the one item scale *General Presence*. Each subscale ranges from 0 - 6.

To consider the player's physical activity as a potential moderator for the assessment of the exergame and the player experience, we used the International Physical Activity Questionnaire (IPAQ) [13] to assess the player's physical activity level. There are two forms of output, a categorization of the participant's activity level (low, moderate, high activity) and a total value representing the weekly amount of energy expenditure carrying out physical activity. Furthermore, we used the Simulator Sickness Questionnaire (SSQ) [23] to assess if our setup causes simulator sickness. The resulting SSQ score can be between 0 and 235.62.

B. Design and Procedure

The study was conducted in a between-subjects design. Hence, each participant either played the TV or VR version of the exergame. We chose not to use a within-subjects design to prevent potential exhaustion, boredom, and fatigue of the participants. We deliberately did not include a third condition in which participants would have engaged in physical activity without an exergame, because a comparison between this variant and a non-VR exergame [1], [17] as well as to a VR exergame [26], [35] has already been studied in detail.

After arriving at the study site, the participants put on the pulse belt gave their consent to the study and completed questionnaires on demographics, prior experience with exergames and VR, IPAQ as well as the SSQ to check whether any high SSQ values after gaming were due to high SSQ values before gaming. During that starting period, we recorded the participants' heart rate to obtain a baseline heart rate, which we then could compare to the heart rate during the gameplay. The starting period depended on how fast the participant filled out the questionnaires but took around 7 minutes on average. Afterward, the gameplay phase started. First, the participants could then choose which hand they wanted to use for each controller. In the TV condition, the participants were positioned about 2 meters away from the screen. In the VR condition, they put on the HMD and were positioned at the center of a 3m * 2m play area. After finishing a tutorial on the basic gameplay, the participants started the game themselves. The gameplay in each condition lasted five minutes. We deliberately decided on a duration that falls between the duration of the works of [3] and [17], both of which examined arousal in exergames. During the gameplay phase, we recorded the participants' heart rates. Immediately after completing the game, the participants first completed the SAM and Stroop test. Subsequently, the SSQ, Borg, IMI, and IPQ were completed. The study ended with the participants being debriefed about the study. The study was compensated with one subject hour, of which up to 30 must be collected in the context of different study programs at our university. Since the Stroop test is based on colors, only participants without color vision deficiencies were allowed to participate in the evaluation. The ethics committee of our university approved our study. As the study took place during the COVID-19 pandemic, a special hygiene concept was developed and accepted by our university.

C. Hypotheses

Based on the assumption that VR applications can elicit higher arousal than non VR applications [10], [15], [32], we hypothesize that:

I *The VR condition leads to a significantly increased arousal compared to the TV condition.*

Since arousal can have an impact on cognitive performance [24] but can improve or worsen it depending on the height of the arousal level [42], we further hypothesize that:

II *The VR condition leads to a significantly different cognitive performance compared to the TV condition.*

Furthermore, due to the identical gameplay in both variants, we assume that:

III *The exertion in the VR condition is not significantly different from the TV condition.*

Based on the results of [6], we further assume that:

IV *The VR condition leads to a significantly increased experience of presence compared to the TV condition.*

V *The VR condition leads to a significantly increased motivation compared to the TV condition.*

D. Results

20 female and 12 male participants were included in the analysis ($N = 32$) and randomly assigned to one of the two conditions. Both conditions were assigned 16 participants each. The age ranges from $18 - 34$ years ($M = 22.6, SD = 3.12$). 59.38% of the participants reported having played an exergame at least once. 87.5% of the participants in the VR condition had no or just a little experience with VR systems, while the remaining 12.5% had at least five VR experiences. The IPAQ evaluation showed that the physical activity level of 6 participants can be rated as low, 13 as moderate, and 13 as high.

TABLE I
MEANS AND STANDARD DEVIATIONS OF THE DEPENDENT VARIABLES WHICH ARE NOT PRESENTED IN FIGURES 3 AND 4

Subscales	VR M *(SD)*	TV M *(SD)*
IPQ		
Spatial Presence	4.33 (*0.95*)	1.93 (*1.34*)
Involvement	3.39 (*1.47*)	2.16 (*1.19*)
Experienced Realism	1.78 (*0.93*)	0.97 (*0.72*)
IMI		
Perceived Competence	4.10 (*1.16*)	2.58 (*1.28*)
Pressure/Tension	4.26 (*1.25*)	4.54 (*1.35*)
SAM		
Valence	7.26 (*1.61*)	5.25 (*1.88*)
Dominance	5.56 (*1.71*)	3.38 (*1.31*)
Stroop		
Incongruent Stimuli Response Time	27.6 (*6.08*)	26.7 (*6.04*)
Congruent Stimuli Response Time	8.64 (*2.42*)	9.33 (*2.97*)

We calculated independent t-tests for the following group comparisons if normal distribution and variance homogeneity was given. If normal distribution was given, but no variance homogeneity was given, we calculated Welch t-tests. For both variants of the t-test, we used Cohen's d as the effect size measure. According to [12] the values of $|d| = 0.2$, 0.5, and 0.8 correspond to small, medium, and large effect sizes. If no normal distribution was given, we calculated Mann-Whitney-U tests and used the coefficient of determination R^2 as effect size measures. Values of $R^2 = 0.02$, 0.13, and 0.26 correspond to small, medium, and large effect sizes [12]. We calculated a mixed ANOVA with the condition as between factor and time of measurement as within factor for the heart rate. Normal distribution, equality of covariance matrices, and variance homogeneity were given. For the mixed ANOVA, we used eta-squared η^2 as the effect size measure. Values of $\eta^2 = 0.02$, 0.06, and 0.14 correspond to small, medium, and large effect sizes [12]. The variables relevant to the hypotheses are shown in figures 3 and 4. The remaining descriptive values, which are not represented in the figures, are shown in table I. The inferential statistical comparisons are presented in the following subsections.

1) ***Presence****:* A Mann-Whitney-U test revealed significantly higher *General Presence* values in the VR condition $U = 17.5, Z = -4.25, p < .001, R^2 = 0.56$. Further, t-tests revealed significantly higher values in the VR condition for the *Spatial Presence* $t(30) = 5.85, p < .001, d = 2.07$, *Involvement* $t(30) = 2.67, p = .012, d = 0.95$, and *Experienced Realism* $t(30) = 2.77, p = .010, d = 0.98$ subscale.

2) ***Motivation****:* A Mann-Whitney-U test revealed significantly higher *Interest/Enjoyment* values in the VR condition $U = 66.5, Z = -2.32, p < .020, R^2 = 0.17$. For the *Perceived/Competence* subscale, a t-test revealed significantly higher values in the VR condition $t(30) = 3.53, p = .001, d = 1.25$. No significant difference was found for the

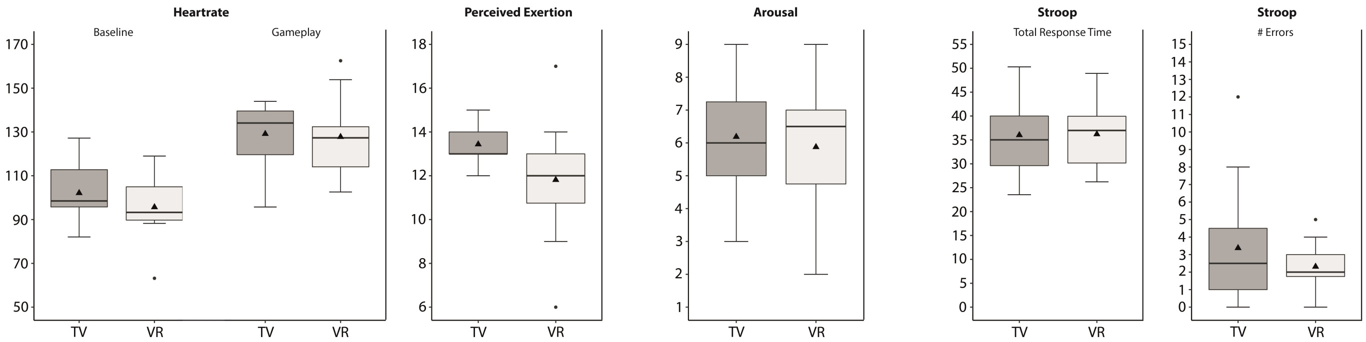

Fig. 3. Boxplots of the constructs relevant to hypotheses I-III.

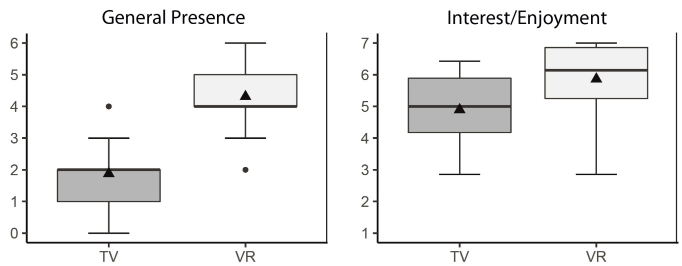

Fig. 4. Boxplots of the constructs relevant to hypotheses IV&V.

Pressure/Tension subscale $t(30) = -.60, p = .554, d = -.21$.

3) ***Exertion:*** For the Borg Scale, a t-test revealed significantly lower values in the VR condition $t(18.98) = -2.51, p = .021, d = -0.89$. For the heart rate we had to exclude nine participants since the pulse belt did not record data for these cases. A mixed ANOVA revealed no significant interaction between condition and time of measurement $F(1, 21) = 0.49, p = .490, \eta^2 = .004$. There was a significant main effect for time of measurement $F(1, 21) = 66.45, p < .001, \eta^2 = .481$ but no significant main effect for condition $F(1, 21) = 0.50, p = .487, \eta^2 = .008$.

4) ***Arousal:*** Mann-Whitney-U tests revealed significantly higher values in the VR condition for the *Dominance* $U = 39.5, Z = -3.39, p = .001, R^2 = 0.36$ and *Valence* $U = 43.05, Z = -3.28, p = .001, R^2 = 0.34$ subscale. However, a t-test could not reveal a significant difference for the *Arousal* subscale $t(30) = 3.53, p = .645, d = -.17$.

5) ***Cognitive Performance:*** With the exception of the Number of Errors, the following tests apply to all answers, regardless of whether they were correct or incorrect. Neither t-tests for the total response time $t(30) = 0.07, p < .941, d = 0.03$ and the response time of the incongruent stimuli $t(30) = 0.41, p < .684, d = 0.15$, nor Mann-Whitney-U tests for the response time of the congruent stimuli $U = 114.0, Z = -0.53, p = .598, R^2 < 0.01$ and the number of errors $U = 117.0, Z = -0.42, p = .674, R^2 < 0.01$ revealed significant differences between the conditions.

6) ***SSQ, IPAQ, and Correlations:*** The SSQ values in the VR condition before $M = 25.01, SD = 30.50$ and after $M = 29.20, SD = 36.72$ the gameplay are very low and give no reason to believe that simulator sickness has had a confounding influence. The IPAQ values do not significantly differ between the conditions $t(30) = -0.797, p = .431, d = -0.28$. Interestingly, we have found no correlations between the Stroop constructs and any of the other assessed subscales. However, beyond the cognitive performance, we have found correlations between the *Interest/Enjoyment* subscale and the *Spatial Presence* subscale $r(30) = .364, p = .041$, between the *Interest/Enjoyment* subscale and the *Borg* value $r(30) = -.363, p = .041$, and between the *Spatial Presence* subscale and the *Borg* value $r(30) = -.434, p = .013$.

V. Discussion

Our evaluation was motivated by whether cognitive performance following an exergame can be improved when the exergame is played in VR instead of in front of a TV screen. We considered arousal as a central moderator for the change in cognitive performance. We assumed that the use of VR could significantly increase arousal. We observed significant differences for the emotional dimensions valence and dominance, but not for arousal. Thus, we reject our first hypothesis. Hence, our results are not in line with [10], [15], [32], who found increased arousal levels in VR applications in contrast to non-VR applications. This leads to the question of why we could not observe this difference in our results. A possible explanation for this could be that the exergame used in our study already elicited arousal to an extent where the additional use of VR could not or only insignificantly increase the arousal level further. Moreover, the nature of the visual stimuli is crucial when it comes to the increase of arousal by VR systems [15]. Thus, our virtual environment's design and visual appeal could have just been perceived as not arousing. However, our VR exergame could significantly increase the experienced valence and dominance. Thus, we conclude that VR can substantially impact the player's affective response, but in the context of exergames, it has no impact on the player's arousal. Based on the results of [17], we, therefore, assume arousal is mainly influenced by the exergame regardless of the output device.

Similar to the arousal scores, the evaluation of the Stroop test reveals no significant difference between the conditions.

We thus also reject our second hypothesis. Hence, according to our results, the player's cognitive performance after a one-time exergaming intervention does not seem to benefit from the use of VR. Thus, VR exergaming interventions that aim to increase the player's cognitive performance, such as in [35], mainly profit from the exergame and not from the VR system. As in [1], the exergame itself seems to have the most significant impact on cognitive performance, regardless of the output medium. Interestingly, our results are congruent to those of [21], although the game used, the target audience, the VR system, and the duration of the game intervention differ. Interestingly, however, the authors of [21] observed a positive impact of VR on cognitive performance after a 4-week intervention. Thus, although we conclude that VR can not improve cognitive performance after a single exergame intervention, VR could be beneficial for long-term exergame interventions that aim to support cognitive performance.

Besides the main question regarding arousal and cognitive performance, we have assessed the players' exertion, motivation, and presence to give a more detailed insight into the effects of VR on the player experience. The evaluation of the heart rate shows, as hypothesized, no significant difference between the conditions. It is noticeable that the baseline heart rate is relatively high in both conditions. This could be because we did not have a resting phase at the beginning of the evaluation in which the baseline was measured since we were not interested in the exact resting heart rate but instead in checking if the heart rate change in both conditions was similar. Interestingly, the evaluation of perceived exertion shows significantly lower values in the VR condition, although physiological exertion does not differ between conditions. Thus, we only partly confirm our third hypothesis. We assumed that, as in [6], perceived exertion shows the same tendency as actual exertion since our game does not require repetitive interactions such as rowing or cycling, which could be performed without a constant focus. Instead, the VR condition appears to have distracted from exertion as in the work of [28]. Since we have found negative correlations between the presence/motivation and the perceived exertion, our results suggest that also for VR exergames with constantly changing interactions, VR and the resulting presence and motivation can decrease the perceived exertion.

Furthermore, the evaluation of the IPQ reveals significantly higher values for all subscales in the VR condition than in the TV condition. Thus, the VR condition of the exergame led to a higher sense of presence. Our assumption that the VR system is more immersive than the TV system and thus leads to a higher presence is thus confirmed. Similar to the results of [6] the output device seems to have a significant impact on the experience of presence. We, therefore, accept hypothesis IV.

The evaluation of the IMI reveals higher scores for the *Interest/Enjoyment* and *Perceived Competence* subscales in the VR condition. Thus, the VR version of the exergame could lead to significantly higher intrinsic motivation than the TV condition. Hence, we also accept hypothesis V. Our results are in line with the findings of [6], [28] and confirm the positive influence of VR on players' intrinsic motivation in the context of exergames. Of particular interest is also the positive correlation between the Interest/Enjoyment subscale and the Spatial Presence subscale, which, in the context of exergames, was already observed by [7]. Since intrinsic motivation is the main object of investigation in most work on (VR) exergames, future research and developments should pay special attention to a potential connection between the players' presence and motivation.

A limitation of our work is that we only surveyed 32 participants, which could result in reduced test power. Also, our sample has few participants with a low physical activity level, which indicates a high basic motivation for physical activities and could thus have a particular impact on the assessed motivation. Moreover, the participants in our study were predominantly young, which makes it difficult to transfer our results to other age structures, especially in the context of cognitive performance. Thus, future studies examining the impact of VR exergames on cognitive performance should include a more diverse sample of participants regarding their age and physical activity level. Further, future studies should also vary the duration of the game and the type of physical movements, as both factors can have a substantial impact on cognitive performance [24]. Furthermore, future studies should measure more aspects of cognitive performance and use additional physiological parameters to measure arousal.

VI. Conclusion

Our study indicates that using a VR system instead of a conventional TV screen system does not benefit cognitive performance for a one-time exergame intervention. Moreover, our results suggest that while VR can enhance parts of the player's emotional response, arousal, as a vital moderator for cognitive performance, cannot be increased by the mere use of a VR system in the context of exergames. Beyond that, however, our results show that VR can significantly benefit the player's motivation and reduce perceived exertion. Particularly the player's experience of presence resulting from the high immersion the VR system offers seems to be linked with this increase. Our results help future developers as we show the benefits and limitations of VR systems for exergames. Future exergame developments that focus on player motivation should consider VR exergames, while exergames with the goal of increasing cognitive performance do not necessarily need to use VR to achieve a significant benefit.

References

[1] Cay Anderson-Hanley, Paul J Arciero, Adam M Brickman, Joseph P Nimon, Naoko Okuma, Sarah C Westen, Molly E Merz, Brandt D Pence, Jeffrey A Woods, Arthur F Kramer, et al. Exergaming and older

adult cognition: a cluster randomized clinical trial. *American journal of preventive medicine*, 42(2):109–119, 2012.

[2] Sebastian Arndt, Andrew Perkis, and Jan-Niklas Voigt-Antons. Using virtual reality and head-mounted displays to increase performance in rowing workouts. In *Proceedings of the 1st International Workshop on Multimedia Content Analysis in Sports*, pages 45–50, 2018.

[3] Soumya C Barathi, Michael Proulx, Eamonn O'Neill, and Christof Lutteroth. Affect recognition using psychophysiological correlates in high intensity vr exergaming. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2020.

[4] Gillian Barry, Brook Galna, and Lynn Rochester. The role of exergaming in parkinson's disease rehabilitation: a systematic review of the evidence. *Journal of neuroengineering and rehabilitation*, 11(1):33, 2014.

[5] Borg G Borg's. Perceived exertion and pain scales. *Champaign, IL: Human Kinetics*, 1998.

[6] Felix Born, Sophie Abramowski, and Maic Masuch. Exergaming in vr: The impact of immersive embodiment on motivation, performance, and perceived exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pages 1–8. IEEE, 2019.

[7] Felix Born, Maic Masuch, and Antonia Hahn. Ghost sweeper: Using a heavy passive haptic controller to enhance a room-scale vr exergame. In *2020 IEEE Conference on Games (CoG)*, pages 221–228. IEEE, 2020.

[8] Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[9] Margaret M Bradley and Peter J Lang. International affective digitized sounds (iads): Stimuli, instruction manual and affective ratings (tech. rep. no. b-2). *Gainesville, FL: The Center for Research in Psychophysiology, University of Florida*, 1999.

[10] Sandra L Calvert and Siu-Lan Tan. Impact of virtual reality on young adults' physiological arousal and aggressive thoughts: Interaction versus observation. *Journal of applied developmental psychology*, 15(1):125–139, 1994.

[11] Yu-Kai Chang, Jeffrey D Labban, Jennifer I Gapin, and Jennifer L Etnier. The effects of acute exercise on cognitive performance: a meta-analysis. *Brain research*, 1453:87–101, 2012.

[12] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[13] Cora L Craig, Alison L Marshall, Michael Sjorstrom, Adrian E Bauman, Michael L Booth, Barbara E Ainsworth, Michael Pratt, ULF Ekelund, Agneta Yngve, James F Sallis, et al. International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise*, 35(8):1381–1395, 2003.

[14] Edward L Deci and Richard M Ryan. Intrinsic motivation inventory (imi) [www document]. http://selfdeterminationtheory.org/intrinsic-motivation-inventory/, 2003. [Online; accessed 11-September-2018].

[15] Sergio Estupiñán, Francisco Rebelo, Paulo Noriega, Carlos Ferreira, and Emília Duarte. Can virtual reality increase emotional responses (arousal and valence)? a pilot study. In *International conference of design, user experience, and usability*, pages 541–549. Springer, 2014.

[16] Samantha Finkelstein, Andrea Nickel, Zachary Lipps, Tiffany Barnes, Zachary Wartell, and Evan A Suma. Astrojumper: Motivating exercise with an immersive virtual reality exergame. *Presence: Teleoperators and Virtual Environments*, 20(1):78–92, 2011.

[17] Yue Gao and Regan Mandryk. The acute cognitive benefits of casual exergame play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1863–1872, 2012.

[18] Zan Gao and S Chen. Are field-based exergames useful in preventing childhood obesity? a systematic review. *Obesity Reviews*, 15(8):676–691, 2014.

[19] Negin Hamzeheinejad, Daniel Roth, Daniel Götz, Franz Weilbach, and Marc Erich Latoschik. Physiological effectivity and user experience of immersive gait rehabilitation. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1421–1429. IEEE, 2019.

[20] Carrie Heeter. Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, 1(2):262–271, 1992.

[21] Kuo-Ting Huang. Exergaming executive functions: An immersive virtual reality-based cognitive training for adults aged 50 and older. *Cyberpsychology, Behavior, and Social Networking*, 23(3):143–149, 2020.

[22] Christos Ioannou, Patrick Archard, Eamonn O'Neill, and Christof Lutteroth. Virtual performance augmentation in an immersive jump & run exergame. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2019.

[23] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[24] Kate Lambourne and Phillip Tomporowski. The effect of exercise-induced arousal on cognitive task performance: a meta-regression analysis. *Brain research*, 1341:12–24, 2010.

[25] Peter J Lang. International affective picture system (iaps): Affective ratings of pictures and instruction manual. *Technical report*, 2005.

[26] Xiaoxuan Li, Kavous Salehzadeh Niksirat, Shanshan Chen, Dongdong Weng, Sayan Sarcar, and Xiangshi Ren. The impact of a multitasking-based virtual reality motion video game on the cognitive and physical abilities of older adults. *Sustainability*, 12(21):9106, 2020.

[27] Colin M MacLeod. The stroop task: The" gold standard" of attentional measures. *Journal of Experimental Psychology: General*, 121(1):12, 1992.

[28] Daniel J McDonough, Zachary C Pope, Nan Zeng, Wenxi Liu, and Zan Gao. Comparison of college students' blood pressure, perceived exertion, and psychosocial outcomes during virtual reality, exergaming, and traditional exercise: An exploratory study. *Games for Health Journal*, 2020.

[29] Albert Mehrabian. A semantic space for nonverbal behavior. *Journal of consulting and clinical Psychology*, 35(2):248, 1970.

[30] Mohammadhossein Moghimi, Robert Stone, Pia Rotshtein, and Neil Cooke. Influencing human affective responses to dynamic virtual environments. *Presence: Teleoperators and Virtual Environments*, 25(2):81–107, 2016.

[31] Florian'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G Sheridan. Designing sports: a framework for exertion games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2651–2660. ACM, 2011.

[32] Federica Pallavicini, Alessandro Pepe, and Maria Eleonora Minissi. Gaming in virtual reality: what changes in terms of usability, emotional response and sense of presence compared to non-immersive video games? *Simulation & Gaming*, 50(2):136–159, 2019.

[33] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[34] Richard M Ryan, Geoffrey C. Williams, Heather Patrick, and Edward L Deci. Self-determination theory and physical activity: the dynamics of motivation in development and wellness. *Hellenic journal of psychology*, 6:107–124, 2009.

[35] Borja Sañudo, Ellie Abdi, Mario Bernardo-Filho, and Redha Taiar. Aerobic exercise with superimposed virtual reality improves cognitive flexibility and selective attention in young males. *Applied Sciences*, 10(22):8029, 2020.

[36] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001.

[37] Mel Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557, 2009.

[38] Mel Slater and Martin Usoh. Body centred interaction in immersive virtual environments. *Artificial life and virtual reality*, 1(1994):125–148, 1994.

[39] Gijsbert Stoet. Psytoolkit: A software package for programming psychological experiments using linux. *Behavior research methods*, 42(4):1096–1104, 2010.

[40] Gijsbert Stoet. Psytoolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1):24–31, 2017.

[41] Wenge Xu, Hai-Ning Liang, Zeying Zhang, and Nilufar Baghaei. Studying the effect of display type and viewing perspective on user experience in virtual reality exergames. *Games for health journal*, 9(6):405–414, 2020.

[42] Robert M Yerkes, John D Dodson, et al. The relation of strength of stimulus to rapidity of habit-formation. *Punishment: Issues and experiments*, pages 27–41, 1908.

[43] Nan Zeng, Zachary Pope, and Zan Gao. Acute effect of virtual reality exercise bike games on college students' physiological and psychological outcomes. *Cyberpsychology, Behavior, and Social Networking*, 20(7):453–457, 2017.