Effects of Valence and Arousal on Working Memory Performance in Virtual Reality Gaming

Daniel Gabana, Laurissa Tokarchuk
School of Electrical Engineering and Computer Science
Queen Mary University of London, UK
d.gabanaarellano, laurissa.tokarchuk
@qmul.ac.uk

Emily Hannon
School of Biological and Chemical Sciences
Queen Mary University of London, UK
e.hannon@qmul.ac.uk

Hatice Gunes
Computer Laboratory,
University of Cambridge, UK
hatice.gunes@cl.cam.ac.uk

Abstract—The role of affective states in cognitive performance has long been an area of interest in cognitive science. Recent research in game-based cognitive training suggest that cognitive games should incorporate real-time adaptive mechanisms. These adaptive mechanisms would change the game's difficulty according to the player's performance in order to provide appropriate challenges and thus, achieve a real cognitive improvement. However, these mechanisms currently ignore the effects of valence and arousal on the player's cognitive skills. In this paper we investigate how working memory (WM) performance is affected when playing a VR game, and the effects of valence and arousal in this context. To this aim, a custom video game was created for Desktop and VR. Three difficulty levels were designed to evoke different levels of arousal while maintaining the same memory load for each difficulty level. We found an improvement in WM performance when playing in VR compared to Desktop. This effect was particularly pronounced in those with a low WM capacity. Significantly higher levels of valence and arousal were self-reported when playing in VR. We explore the impact that reported affective states could have in the player's WM performance. We suggest that high levels of arousal and positive valence can lead players to a flow state [1] that may have a positive impact on the player's WM performance.

1. Introduction

The influence of affective states in attention and WM have been widely investigated [2] [3] [4]. These two cognitive skills are intrinsically related; attention regulates the incoming information and WM retains it while other cognitive processes are ongoing [5]. Whilst arousal has been demonstrated to enhance attention up to a certain point, after which it has a negative effect, the role of valence is still considered controversial and might be task dependent [2]. Recent research in gaming environments suggest that negative valence does not have consistent effects on WM [2] [4]. However, attention can be increased using appropriate challenges that activate the player [4].

Video games have been used to train cognitive control and treat emotional disorders such as depression [6] [7] [8]. The current approach in recent game-based cognitive training research is to include adaptive mechanisms that changes the game's difficulty in real-time depending on the players' performance [6] [7]. However, these adaptive mechanisms ignore the important role of affective states on cognitive performance. Including the detection and analysis of affective states in these adaptive mechanisms can improve the adaptation and thus, have a positive impact in the player's cognitive performance [9].

The aims of this study are two-fold: 1) investigating WM performance when playing a video game in VR compared to Desktop, and 2) exploring the role of valence and arousal on WM performance.

2. Related work

2.1. Cognition and affect

Affective states, represented in a two dimensional space of valence and arousal [10], can have positive or negative effects on attention and WM. Bennion et al. [2] summarised the main hypotheses about how emotion affects memory. While positive and negative valence usually enhance memory, arousal aids memory up to a point (it has a detrimental influence when it is too high). Furthermore, when emotion helps to process information (encoding), it also facilitates the storage of that particular information (consolidation) [2]. This relationship between affect and cognition has also been studied in physiological terms using mainly Heart Rate Variability (HRV) as a measure of stress and cognitive load. A recent study looking at the influence of three visuo-spatial WM loads on HRV demonstrated that poor WM performers showed lower HRV than good performers [11].

Yeh et al. [4] investigated how negative emotions influence WM in a creativity game asking participants to play three increasing difficulty levels. Similar to Bennion et al. [2], their findings indicated that arousal could help WM up to a point but then had negative effects. They proposed

a game that presents challenges to activate the player's attention while avoiding negative emotions. However, self-reported negative emotions did not have consistent effects on WM, probably due to subjective factors such as motivation [4] [2].

2.2. Affect detection in gaming

The usage of VR in video gaming in the last few years has brought an increase in the player's immersion and engagement. This new degree of immersion, referred to as *presence*, is reported as the feeling of *'being there'* in the virtual world [12]. Research linking presence and affective states has suggested that higher levels of presence directly influence the vividness and intensity of the emotions users experience [12]. Thus, VR has been referred to as an *'affective medium'* due to its ability to evoke and intensify the affective states [12].

Affective states are used for providing adaptive feedback, i.e. closed-loop video games, which consists of adapting the game according to the player's current affective states [13]. The detection of affective states can be used not only to improve the player experience but also for therapeutic purposes [14]. Welch et al. [15] used machine learning algorithms to create a therapeutic VR closed-loop video game for children with autism that adapts to the player's affective states to maintain an optimal engagement level and enhance their social skills.

2.3. Affect in cognitive training

Video games have long been studied for their effects on cognitive skills. However, research has shown that not all games induce cognitive benefits. Studies where subjects repeated trials of a gamified cognitive task without difficulty adaptation only led to faster reaction times [13] [5]. Recently, the video game *Neuroracer* created by Anguera et al. [6] has demonstrated real improvements in older adult's cognitive skills like attention and multitasking. This game adapts the difficulty level to the player's performance in order to appropriately challenge their cognitive skills. Using this approach, other video games like *Project: EVO* have been created to treat cognitive and emotional disorders such as depression [7]. These games add certain gaming elements such as rewards or storyline in order to make the game fun and engaging [13].

This new generation of games for cognitive training demonstrates the importance of intelligent adaptive mechanisms to challenge the player and develop personalised training. According to Mishra et al. [13], the input channel of closed-loop video games should use not only the player's performance metrics but also real-time data from the player's interactions and behaviour. Using physiological or motion sensors to measure the player's affective states can result in a more accurate adaptation loop that could enhance the effects of these new cognitive treatment games.

2.4. Contributions of this work

We investigate the effects of playing a video game in VR to the player's WM performance. We hypothesise that due to the high level of immersion in VR [12], participants would have a better WM performance than playing in a Desktop environment. We also explore the effects that arousal and valence have on WM performance and how they can be used in a closed-loop video games for cognitive training.

3. The study

The aims of this study are: 1) to investigate WM performance when playing a video game in Desktop and VR settings, and 2) to explore the effects of valence and arousal on WM performance. For this purpose, we developed a custom endless running game[1] for Desktop and VR called *Memory Break* inspired by the iOS game *Smash Hit*[2]. This type game of game, similar to *Project: EVO* [7], was chosen due to its simplicity, low cognitive verbal demands and easiness to adapt for both Desktop and VR. Three difficulty levels - easy, medium and hard - were designed to evoke three levels of arousal: low, medium and high arousal respectively. The difficulty levels are exactly the same in Desktop and VR. To keep the interaction as similar as possible, players interact using the hand-gesture tracking device Leap Motion[3] in both settings. While Leap Motion is mounted in front of the HMD in the VR setting, in the Desktop setting it is placed on top of the desk (see Fig. 1). The difference of location impacted how the hand gesture could be tracked and thus the exact gesture interaction used in each mode has some differences.

3.1. The game: *Memory Break*

The goal of the game is to obtain the highest score possible. It consists of throwing balls at different stationary or moving obstacles to successfully pass through without crashing into them. If the player crashes, five points are deducted from the score; one point is also subtracted every time a ball is thrown. In order to get points, the participants have to throw and hit the green gems found on their way, which added 10 or 20 points depending on the type of gem collected.

Each level is divided into five sections structured as follows: Every 30s of game play, the game stops at a door where a random sequence of letters appears that has to be remembered. These letters appear one at a time for 800ms, with 500ms gaps between them. The sequence length is randomly selected between 3 and 7 letters, each sequence only appearing once per level. This results in a total of 25 letters per level, the same number of letters as one block of the Operation Span test [16], a complex memory task that employs maths and letters to assess WM capacity (see 3.3). After the sequence is displayed, the doors open and

1. https://en.wikipedia.org/wiki/Platform_game#Endless_running_game
2. http://www.smashhitgame.com/
3. https://www.leapmotion.com/

Figure 1. Screenshot of *Memory Break* in the VR and Desktop

the game continues for another 7s. The game then stops again at another door where the player has to recall and input the letters previously shown in the exact same order. Once completed, feedback is provided indicating how many letters have been recalled correctly, the doors open and the game continues to the next section.

The game's speed is incremented at each difficulty level by 5%. The number of obstacles between doors is also incremented by 2 units, as well as the average number of obstacles per section, being 8, 17 and 34 for the easy, medium and hard level respectively. The duration of each difficulty level is approximately five minutes.

3.2. Testing the game

We tested the game before the actual study to assess whether the three difficulty levels evoked different levels of arousal. Seven participants took part in a pilot study with a mean age of 28.57. Participants reported a mean arousal level of 0.7 (SD:0.16), 0.73 (SD:0.18) and 0.79 (SD:0.16) in the easy, medium and hard levels. As a result, the easy level was modified further in order to make it easier. Some improvements to the game were made according to the participants' suggestions such as adding audio feedback when touching buttons or changing certain colours.

3.3. Procedure and measurements

The study consisted of two sessions, one for each interaction mode. Participants were asked to leave one week between sessions and book a follow-up session at the same time and day of the week to minimise differences that might result from being tested at different times of the day. During the first session, participants had to fill in a background and gaming habits questionnaire [8], followed by the shortened version of the Automated OSpan test [16] that measured their WM capacity baseline. This test uses maths problems and letters that have to be remembered to assess the WM capacity.

Before introducing the game, a Polar H7 heart rate sensor was attached to the participant's chest and paired

with the HRVLogger [17] app that recorded the subject's heart activity. Once participants relaxed and the Heart Rate (HR) baseline was recorded, the experimenter explained the game mechanics and interaction control. Before playing the game in VR, participants played a 3 minutes VR game called *Blocks*[4] where they could get used to the HMD, the virtual world and the interaction with their virtual hands. This aimed to reduce the novelty effect that a VR experience could cause on affective states [18]. After a practice play with our game *Memory Break*, each participant played each level once in random order. Immediately after playing each difficulty level, participants reported their level of engagement and completed the In-game module of the Experience Questionnaire [19], as well as self-reported the level of arousal and valence using the Affective Slider [20]. Finally, when all difficulty levels were played, participants completed the Post-game module of the Game Experience Questionnaire [19] and rated each levels in terms of difficulty, boredom, enjoyment, arousal and focus (see Fig. 2). In this paper, we analyse 4 out of the 7 components of the In-game module: competence, tension, immersion and flow.

3.4. Participants

Thirty participants, 15 male and 15 female, with mean age of 26.43 (SD: 4.8) were randomly assigned to one of the interaction modes in their first session. None of the selected participants had been diagnosed with any learning difficulty such as dyslexia. 43% of the participants reported to have played video games between 0 and 2 hours the week before the study, while 33% did not play any. All the participants reported that they liked the game overall in both interaction modes, except 3 participants who disliked the game in their second session when playing in Desktop setting. Most of the participants (87%) found the Desktop interaction difficult to manage, and only 27% struggled in VR. This was mainly due to the different location of LM in Desktop and VR (see Fig. 1). The placement of this sensor in front of the HMD made the VR interaction easier and more natural to control, while the Desktop interaction was reported to be less natural and more uncomfortable.

4. Analysis and results

No pre-processing was needed on the HR or HRV features as they were computed by the HRVLogger app [17]. Different HRV features such as the average of normal-to-normal intervals (AVNN), root mean square of successive differences (rMSSD) and the low-high frequency ratio (LFHF) were selected for the analysis. These features were chosen based on previous research linking HRV and visuo-spatial WM performance [11]. Prior to the analysis, each participant's HR and HRV were normalised dividing the mean of each difficulty level by their baseline mean. This normalisation procedure scales the values so that all participants have a common mean of 1. WM capacity and

---

4. https://gallery.leapmotion.com/blocks

WM performance scores were also normalised dividing the number of letters recalled correctly by the total number of letters presented, which results in the percentage of letters recalled correctly.

We carried out repeated measures Analysis of Variances (ANOVA) to assess the statistical differences between interaction modes and difficulty levels. F-values and significance levels (p) are reported. Further post-hoc t-test analyses were conducted where relevant. Spearman's correlations (rho) were used to estimate the relationship between self-reported variables. Linear Mixed Effects models were employed to predict WM performance and level of immersion reported. Likelihood ratio tests were conducted to assess the goodness of fit of our models.

4.1. Interaction modes and difficulty levels

Participants reported VR as the interaction they enjoyed the most, being also the setting where they felt more immersed, aroused and focused. Desktop was reported challenging due to the difficulties in controlling the interaction. As shown in Figure 2, most of the participants felt more bored in level 1 in both interaction modes, whilst the third level was the most difficult one and where they felt more aroused.

Engagement showed a significant difference between difficulty levels (F= 7.25, p<.01) and interaction modes (F=18.91, p<.001), while immersion was only significantly different between interaction modes (F=25.41, p<.001). Moreover, the normalised mean HR was significantly higher in VR compared to Desktop (F=9.70, p<.001). In terms of difficulty levels, level 3 evoked the highest HR (F=24.53, p<.001). Nevertheless, these results could be affected by how much participants had to move their hand and head (in VR) to succeed in each difficulty level. Among the HRV features extracted, only AVNN showed significant differences between difficulty levels (F=35.42, p<.001) and interaction modes (F=8.76, p<.01), while rMSSD did only for difficulty levels (F=8.96, p<.001). The self-reported levels of arousal (F=12.73, p<.001) and valence (F=19.70, p<.001) were statistically lower in Desktop than VR. Furthermore, arousal (rho=0.48, p<.001) and valence (rho=0.56, p<.001) correlated significantly with the reported immersion in both settings, indicating that high levels of immersion lead to an increase in self-reported valence and arousal. These results are in line with previous work measuring affective states in VR environments [12].

4.2. High and low working memory

One of the main interests of this study was to investigate the effects of game playing in VR on WM performance. An ANOVA analysis on each interaction mode separately showed a higher but not significant difference in VR (F=2.20, p=.09) than in Desktop (F=0.52, p=.67). A further post-hoc analysis between the normalised WM baseline and WM performance in each level was undertaken, only showing significant differences in VR's level
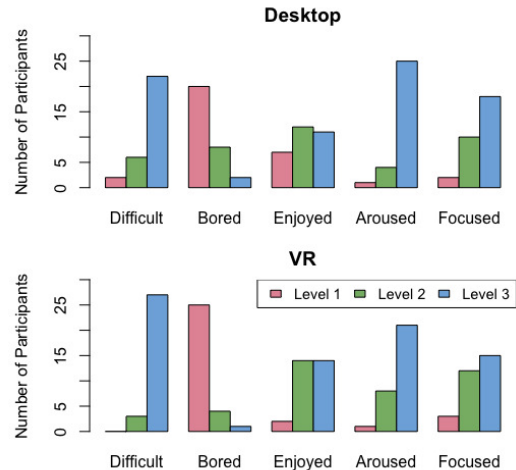

Figure 2. Self-reported comparison levels in each interaction mode

1 (t=-2.18, p<.05) and 2 (t=-2.15, p<.05). Moreover, the normalised WM performance scores presented a weak but significant positive correlation with the self-reported competence (rho=0.23, p<.01), and a negative correlation with tension (rho=-0.28, p<.001).

Based on the normalised WM capacity baseline measured with the OSpan [16] test at the beginning of the study, participants were divided into two groups: low and high WM. Those with a lower normalised WM baseline than the overall median (0.83) were assigned to the low WM group, otherwise to the high WM group. This resulted in 13 subjects in the low WM group and 17 in the high WM. A repeated measures ANOVA was performed for each group, only showing significant differences in the low WM group for the difficulty levels (F=4.96, p<.01). A post-hoc paired t-test analysis assessed statistical differences for the difficulty levels in the low WM group compared to their WM capacity baseline. Participants with low WM achieved significantly better WM scores in VR (Fig. 3) in levels 1 (t=-2.18, p<.05) and 2 (t=-3.22, p<.05). Significant differences using two-samples t-tests were also found in the HRV features extracted between these groups. Similar to previous research reported in a visuo-spatial WM task [11], the LFHF ratio (t=4.31, p<.001) was significantly higher for the low WM group.

4.3. Effects of valence and arousal on WM

In this analysis, we primarily focused on the outcomes obtained from Figure 3, analysing the self-reported levels of valence and arousal of the low and high WM groups. As we see in Figure 4, arousal levels increased for all participants in all difficulty levels. According to Bennion et al. [2], arousal has beneficial effects on WM up to a certain point, after which it has a negative effect. Looking at the self-reported arousal of the two groups, the highest level of arousal is observed in the third and most difficult level of

both interaction modes, which correspond to the lowest WM scores, specially for the high WM group in Desktop setting.

The levels of valence showed more interesting results, correlating significantly with WM performance (rho=0.19, p<.01), being particularly pronounced for the high WM group (rho=0.39, p<.001). This indicates that high levels of positive valence improved WM performance. We observed that when valence and arousal are both high, i.e. when participants were challenged but feeling successful,
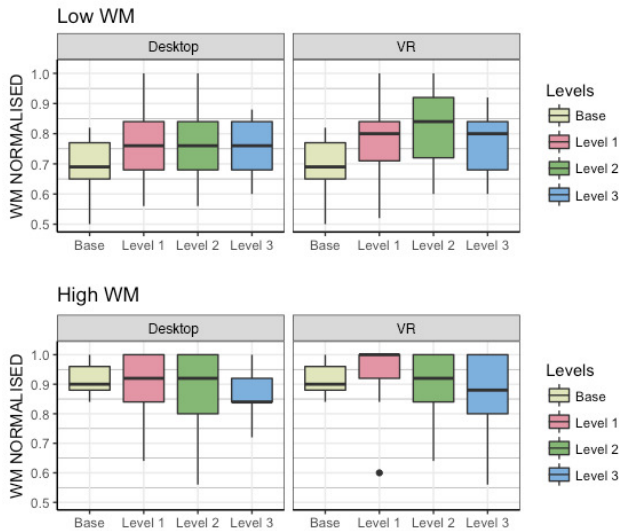

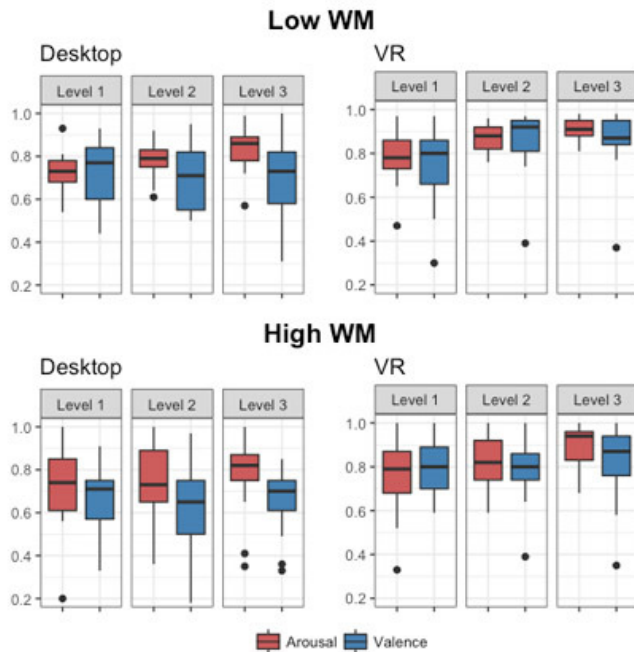Figure 3. WM performance of subjects with low and high WM capacity

Figure 4. Self-reported arousal and valence of players with low and high WM. Note: Arousal and valence are two separate dimensions of affective states; therefore, they are not directly comparable.

they obtained their best WM score. More specifically, WM performance can improve when the player is in a state of enjoyment or flow, described by Csikszentmihályi as the *'optimal experience'* [1]. However, the performance of the low WM group in Desktop's level 1 can be explained by the order of levels played, since half of the low WM participants played level 1 last, having had more experience playing the game and being more relaxed.

Additionally, we used Linear Mixed Effects regression models to predict WM performance and the immersion of participants. We developed a model using the self-reported valence and arousal as fixed and subjects as random effects to predict WM performance. Since we are interested in the relationship between WM and affective states, arousal and valence were the only predictors for the WM model. As seen in Table 1, valence and arousal were significant predictors of WM performance. A likelihood ratio test was performed, showing the goodness of fit of this model (Chi-sq=27.17, p<.001). Using the same random effects, the model to predict immersion used WM, the interaction mode and HRV rMSSD as fixed effects. The inputs of this model explored the relationship between WM, HRV and immersion. The interaction mode and HRV rMSSD were significant predictors of immersion, but not WM, although it improved the model (see Table 2). The likelihood ratio test was also significant for this model (Chi-sq=46.92, p<.001).

5. Discussion

The results presented in this paper show evidence of how game playing in VR can improve WM performance and intensify the self-reported affective states of the players. The normalised mean HR as well as the self-reported levels of arousal and valence were significantly higher in the VR setting. While being immersed and engaged in a VR game, players make better use of their cognitive resources as they are more activated and motivated. The high levels of arousal and valence reported in this interaction mode had a positive effect on the player's cognitive performance, possibly enhancing the capture and encoding of information. This effect was stronger on those with a lower WM capacity, who showed a significant improvement in their WM performance when playing the easy and medium difficulty levels in VR.

TABLE 1. LME MODEL OF WM

	Estimate	Std. Error	df	t	p
(Intercept)	19.54	1.17	175	16.71	<.001
Valence	6.13	1.12	161	5.47	<.001
Arousal	-3.79	1.52	170	-2.49	<.01

TABLE 2. LME MODEL OF IMMERSION

	Estimate	Std. Error	df	t	p
(Intercept)	1.98	.43	170	4.55	<.001
WM	.03	.02	179	1.55	.12
IM (VR)	.62	.09	151	6.66	<.001
HRV rMSSD	-3.79	.19	179	1.88	.06

These results can be linked to the flow theory proposed in [1] known as a state of full immersion and engagement triggered when a subject's skills can overcome challenges, creating a positive experience and positive affect. When players are in this flow state, highly focused and enjoying the game, they make an optimal use of their cognitive skills and thus, a better WM performance is achieved.

One of the limitations of this study was the difference in placement of the hand motion tracking device that led to different gestures in Desktop and VR settings. The reported difficulties in controlling the Desktop interaction might have affected the level of immersion, even though the majority of the participants liked the game in this setting. It is also important to mention that different results could have been obtained using other types of games requiring higher cognitive demands.

6. Conclusions and future work

In this study, we investigated the effects of game playing in Desktop and VR settings on WM performance, as well as the effects of the self-reported arousal and valence on WM performance. A custom video game with a hand-gesture based interaction was developed for both Desktop and VR. Three difficulty levels were created to induce different levels of arousal, maintaining the same memory load for all levels. Physiological and self-reported measures of valence and arousal among other variables such as immersion and engagement were collected.

We conclude that higher levels of self-reported immersion while playing *Memory Break* in VR have a positive effect on the player's WM performance. This improvement was particularly pronounced in participants with low WM capacity, as measured with the automated OSpan test [16]. We also presented suggestions of how self-reported affective states can be beneficial for WM when playing a video game. High levels of arousal and positive valence can create a positive experience, leading players to a flow state [1] that may have a positive impact on the player's WM performance.

We propose to work towards a closed-loop video game in VR that includes the player's affective states in the adaptive loop in order to improve the adaptation. The ideal affective video game for cognitive training should keep the player in an optimal affective state while challenging his or her cognitive skills. Since VR is known to increase the level of immersion [12], a VR game can potentially help players to achieve a better WM performance. Our future work will focus on implementing a machine learning algorithm in our VR game to detect the player's arousal and valence in real-time, in order to automatically adapt the difficulty level to improve the player's WM performance.

Acknowledgments

References

[1] M. Csikszentmihalyi, *Toward a psychology of optimal experience.* Springer, 2014.

[2] K. A. Bennion *et al.*, "Oversimplification in the study of emotional memory," *Journal of the International Neuropsychological Society*, vol. 19, no. 09, pp. 953–961, 2013.

[3] A. Curci *et al.*, "Negative emotional experiences arouse rumination and affect working memory capacity." *Emotion*, vol. 13, no. 5, p. 867, 2013.

[4] Y. Yeh *et al.*, "How stress influences creativity in game-based situations: Analysis of stress hormones, negative emotions, and working memory," *Computers & Education*, vol. 81, pp. 143–153, 2015.

[5] Z. Shipstead *et al.*, "Is working memory training effective?" *Psychological bulletin*, vol. 138, no. 4, p. 628, 2012.

[6] J. A. Anguera *et al.*, "Video game training enhances cognitive control in older adults," *Nature*, vol. 501, no. 7465, pp. 97–101, 2013.

[7] J. A. Anguera *et al.*, "Improving late life depression and cognitive control through the use of therapeutic video game technology: A proof-of-concept randomized trial," *Depression and Anxiety*, pp. 143–153, 2016.

[8] J. A. Anguera *et al.*, "A pilot study to determine the feasibility of enhancing cognitive abilities in children with sensory processing dysfunction," *PloS one*, vol. 12, no. 4, 2017.

[9] J. Mishra *et al.*, "How to assess gaming-induced benefits on attention and working memory," *Games for Health*, vol. 1, no. 3, pp. 192–198, 2012.

[10] H. Gunes *et al.*, "Categorical and dimensional affect analysis in continuous input: Current trends and future directions," *Image and Vision Computing*, vol. 31, no. 2, pp. 120–136, 2013.

[11] M. Suriya-Prakash *et al.*, "Is heart rate variability related to cognitive performance in visuospatial working memory?" *PeerJ PrePrints*, vol. 3, 2015.

[12] G. Riva *et al.*, "Affective interactions using virtual reality: the link between presence and emotions," *CyberPsychology & Behavior*, vol. 10, pp. 45–56, 2007.

[13] J. Mishra *et al.*, "Video games for neuro-cognitive optimization," *Neuron*, vol. 90, no. 2, pp. 214–218, 2016.

[14] J. Gonzalez-Sanchez *et al.*, "Affect measurement: A roadmap through approaches, technologies, and data analysis," 2017.

[15] K. Welch *et al.*, "An affect-sensitive social interaction paradigm utilizing virtual reality environments for autism intervention," *HCI. Ambient, Ubiquitous and Intelligent Interaction*, pp. 703–712, 2009.

[16] J. L. Foster *et al.*, "Shortened complex span tasks can reliably measure working memory capacity," *Memory & cognition*, vol. 43, no. 2, pp. 226–236, 2015.

[17] M. Altini *et al.*, "Hrv4training: Large-scale longitudinal training load analysis in unconstrained free-living settings using a smartphone application," in *Proc. of EMBC 2016*. IEEE, 2016, pp. 2610–2613.

[18] M. Slater *et al.*, "Analysis of physiological responses to a social situation in an immersive virtual environment," *Presence: Teleoperators & Virtual Environments*, vol. 15, no. 5, pp. 553–569, 2006.

[19] W. IJsselsteijn *et al.*, "The game experience questionnaire: Development of a self-report measure to assess the psychological impact of digital games." 2013.

[20] A. Betella *et al.*, "The affective slider: A digital self-assessment scale for the measurement of human emotions," *PloS one*, vol. 11, no. 2, 2016.