Evaluation of a Virtual Gaming Environment Designed to Access Emotional Reactions While Playing

Sarah-Christin Freytag
Department of Psychology and Ergonomics
Technische Universität Berlin
Berlin, Germany
sarah.freytag@mms.tu-berlin.de

Carolin Wienrich
Inst. Human-Computer-Media
Julius-Maximilians Universität Würzburg
Würzburg, Germany
carolin.wienrich@uni-wuerzburg.de

***Abstract*— a virtual environment was designed and implemented in order to examine emotional responses while playing in a virtual reality. A game based design approach was used to create a convincing narrative and meaningful tasks that are simultaneously used to access the users' emotional state. Self-reported ratings on emotions and game experience were incorporated into a background story and conducted while playing. An exploratory study was conducted to evaluate the concept regarding game experience, level of immersion and persuasiveness of the environment. Additionally, differences in emotional state and game experience while playing using the head-mounted display versus using the monitor were assessed. The outcome of individually gathered qualitative data was analyzed. The environment proved suitable for detecting emotional states and game experience and was generally highly accepted by the players. Users reported a high level of presence while engaging with the environment and did not feel a break of immersion in the game as they rated their experience. Further applications, current limitations and possible improvements are discussed.**

Keywords—Virtual Reality; emotion; game flow; game experience; evaluation; user state detection; fear; HMD

I. INTRODUCTION

While head-mounted displays (HMDs) for the use in virtual reality (VR) have only recently primarily become a tool for developers and enthusiasts, low-cost and user-friendly consumer versions have become an accessible extension of the modern gamer's home equipment. With the rise of HMDs usage for entertainment purposes, early strategies to adapt the device to existing content mainly consisted of running games that were created to be displayed via monitor without much adaptation directly on the HMD [1]. Among other problems, one peculiarity became evident: content, which if presented via monitor evoked mild emotional reactions, now caused players to experience those feelings in an emphasized manner. As [2] stated: "*We want you to feel engaged, but that emotional resonance, you kind of get if for free in VR. If anything, I've found that there's a little too much. The focus has to shift how to manage that emotion and how you use it*".

This observation carries a potential impact on not only the guidelines and best practices in creating content for entertainment purposes, but for all occasions in which VR can be used. This includes therapeutic settings [3], educational purposes [4], marketing approaches as well as displays in public spaces.

This created the necessity to generate an approach to access emotional reactions of users during their engagement in VR. Furthermore, if the same content was capable of evoking different reactions if presented via HMD versus via monitor, a means to compare both reactions is needed. While tools that provide post-exposure ratings such as the Game Experience Questionnaire (GEQ) [5] allow for a summarized assessment of the users experience, the shifts and variations in emotional states while interacting in the VR environment remain undocumented. To detect and map those continuous changes, an in-game questionnaire could be used, but would carry the risk of influencing the gaming experience by disrupting the user, lowering their level of engagement and creating an awareness of the experimental setting. If the users' state was to be assessed via behavioral measures only, potentially valuable insight into the users' thoughts and reactions would still remain undocumented.

If, however, it was possible to integrate direct rating measures as well as behavioral data into a complete and persuasive gaming experience, the information lost could be compensated and valuable data can be gathered. Presence is defined as the feeling of "being in the world" [6] or, in relation to VR, of being "there" [7]. To create a natural gaming experience, the user should feel present in a virtual environment. A convincing design of an experience promotes presence [8]. In order to create such a persuasive environment for scientific purposes, a game-based design approach is needed. One of the key aspects is to give the players meaningful tasks and to embed them into a convincing narrative [9]. Furthermore, if the in-game assessment and behavioral data were integrated into such a narrative and respective tasks, the need for presence would be fulfilled while still creating a standardized experimental setup for scientific purposes.

This paper presents the design and implementation of a virtual environment which integrates the aforementioned requirements. Specifically (a) a game-based context with

978-1-5090-5812-9/17/$31.00 ©2017 IEEE

meaningful tasks (b) an in-game questionnaire and behavioral data collection concept integrated into the background story, (c) a means of gradually manipulating a potentially fear evoking stimulus and (d) an option to compare reactions to the same content as displayed in the HMD and monitor respectively. Further, results of a preliminary study evaluating the created environment regarding game experience and persuasiveness are presented and discussed. Results of the comparison of emotions, specifically fear, while playing via monitor or HMD as well as different stimulus conditions, are only briefly introduced and discussed in more detail in [10]. An emphasis lies on introducing results from a qualitative evaluation of the environment and on discussing further improvements, implications and potential applications for this approach.

II. Methods

A. Creation of the Virtual Gaming Environment

1) Conception of the Virtual environment

First, a literature review was conducted on both scientific and industry sources in order to choose an appropriate stimulus as the core element of the virtual environment. Stimuli such as narrow spaces [11], height [12] or degree of realism [11] were compared regarding feasibility and potential to be included into a convincing narrative. Among those factors, "height" was chosen because of its' direct translatability into real world-contexts as well as metric measurability.

Second, using Unity3D, a map was created on which high-rise buildings were distributed in a rectangular layout. A track connecting rooftops was designed (compare fig. 1). Three height variants (v1:15 to v2:18m, 25 to 28m and v3: 45 to 48m) were implemented and pre-selected during a pre-study conducted with four participants (2m , 2f) aged 27 to 30 (*M* 28,75). All participants rated the combination v1 and v3 as most fitting, while giving consistent height ratings for all three variants. Thus, all further development was based on v1 and v3. The track was grouped into three subsections with comparable tasks. Tasks were created in a way that required the user to interact with height as the potentially fear evoking stimulus (A3).

2) In-game user state assessment

Information on the users' state was divided into two areas of interest: First, their emotional state, second, their level of presence and involvement with the game environment during testing. To access anxiety, euphoria, tension, boredom and perceived safety, estimates on a five-point Likert scale were given. The items were adapted for the game context from the short scale for flow of assessment according to [5]. To control differences in height and distance perception, metric estimates of the current height above ground were prompted.

As [13] has discussed, adverse reactions due to a phobic level of fear reflect in stimulus avoidance. In order to gain insight on the magnitude of anxiety and fear, a task that can (a) be easily tracked and quantified, (b) be incorporated as a side quest towards the main goal and (c) be plausibly integrated into the narrative was needed. To account for that, an object-pushing task was introduced. Objects were to be pushed from ledges and lips and tracked via a unique ID. The number and area of the pushed objects was included as a dependent variable.

3) Combining narrative and requirements into one level

In referral to [9], meaningfulness plays a vital part in creating convincing environments. Therefore, a background story that emphasizes the role of the player as an agent with a meaningful task was developed. All communication and assessment requirements were woven into the game story. In particular, it was claimed that, due to their great achievements in sports, the user was invited by a company focusing on building environments for extreme sports to test their creation.

Figure 1 Example of an exposed steel grid pathway connecting two buildings with several objects that can be pushed.

The players' main task was to give ratings on their psychological and physiological state after each segment, in order to see if the track was safe to be opened to the general public. However, the area was described as not completely finished, which is the reason why several trash bags were still strewn about the location. The player was additionally asked to help with the clean-up by pushing all bags they could find on the track into the depth below (thus implementing the object-pushing task described in A2)

Figure 2 Tutorial section with training gantry and a non-activated information point.

The in-game assessment was implemented via three checkpoints. According to the background story, the assistant of the company could be contacted for a short interview after each section as part of the completion of the

main task. A short tutorial section allowed the player to get accustomed to the game play and controls (see fig. 2). Guidance was given by the linear design of the level as well as additional information checkpoints within sight of each other.

B. Post-game state assessment

Additional to the in-game assessment, further measures that were of interest but either too extensive to be integrated into the in-game assessment or were likely to draw the users' attention, were included in a post-game questionnaire.

Simulator sickness and fear share common symptoms [14, 15]. In order to account for misattributions of symptoms evoked by simulator sickness then to that of emotional states, biological symptoms needed to be tracked, as well. The nausea-subscale of the Simulator Sickness Questionnaire (SSQ, [16]) was integrated into the post-game assessment in order to prevent the user from focusing on possible physical symptoms instead of the environment. The core module of the Game Experience Questionnaire (GEQ, [5]) was used to assess overall game flow and involvement. This was especially relevant for gathering information on possible disruptions and breaks of presence during interaction.

The third part of the state assessment was a structured interview aimed at evaluating the virtual environment and the experience. Questions focused on presence, persuasiveness of the virtual environment, controls, game content and the experimental setup. The participants were also given the opportunity to add their own suggestions for improvements and to freely express any ideas or comments regarding their experience.

C. Participants

Eight participants (5m, 3f) aged 27 to 33 years (M 28,1) were pre-screened as active players of video games (14 to 22 years of experience, M 18,3). Only participants with no self-reported susceptibility to motion sickness, fear of heights, epilepsy and claustrophobia were invited.

D. Experimental Setup

The experiment was conducted as within-subject design for presentation mode with four participants each starting with the HMD first, and four starting in the monitor condition. Height was varied between-subjects with four subjects playing the low level condition and four playing the high condition. The monitor was a 24''-model with a resolution of 1980 x 1200 px. The HMD was an Oculus Rift DK2 with a horizontal Field of View of 106.1°. Sound was presented via headphones and the game was controlled with an Xbox360 controller.

E. Procedure

All participants were greeted and given a "mission briefing", giving details on their tasks. They proceeded to play either via HMD or monitor first. At checkpoints, the experimenter, in their in-game role as corporate assistant, communicated with the player and noted their ratings. After finishing the first condition, the post-game questionnaire was filled out. Participants then proceeded to play the next run in their second mode of presentation. Subsequently, the post-game questionnaire was completed again and the structured interview was conducted.

III. Results

Results were grouped according to the two goals of the study.

A. Assessment of emotional reactions

Regarding emotional reactions, intra-individual differences between the in-game ratings from the first checkpoint of each presentation condition were calculated. Averaged differences were grouped by height and compared between the HMD and monitor condition via a directed two-sample t-test. The results showed a significantly higher level of emotional responses (tension ($t(7)= 2.83$, $p= .013$, $d= .999$, euphoria for playing order HMD→monitor ($t(3) = 3$, $p= .029$, $d= 1.5$),) and tendencies (fear, $t(7)=1.528$ $p= .085$, $d= .54$) while playing in the HMD condition. To interpret behavioral data, intra-individual differences for thrown items were calculated for each participant and each condition. A comparison between HMD and monitor condition yielded no significant differences in number of thrown items. Those findings are discussed in-depth in [10].

B. Evaluation of the virtual environment

In respect of the evaluation, the "flow" subscale from the in-game assessment was calculated for each participant individually according to the scoring manual. Averages were calculated for each condition. The unweighted scores of the SSQ "nausea" subscale were calculated according to the manual. The open answers were sighted, tagged and due to their uniqueness, individually estimated.

1) GEQ differences

The averaged GEQ flow scores showed a medium level of game flow during engagement in all four conditions (M 9.5, SD 5.2). This rating represents an intermediate level of experienced game flow that varied little between conditions.

2) Simulator Sickness

An overall high level of simulator sickness was reported among users (depending on height and presentation mode from 16.75(SD 4.92) to 18(SD 6.68)).

3) Qualitative data

Overall, participants showed a positive attitude towards the experience and uniformly expressed experiencing fun during play. Positive aspects were freely named as follows: The *layout* of the level was noted positively. Users felt motivated to further explore the track by seeing the next sections in a distance. Participants regarded the tutorial section as helpful. Single *elements of the track,* specifically a ledge from which the user had to

jump into boxes at the end of the last section were mentioned as being motivating, "fun and challenging at the same time". Throwing objects as *side task* has been perceived as fitting into the overall narrative. However, most participants felt challenged to throw all objects they could find, as individual scores were shown at the end of the track. *The mission* was described as convincing and participants felt good about their character and the reason they were chosen as a test runner. The gaming *environment* as well as the *quality of visuals* were also noted positively and players felt invited to spend time in the environment in order to explore its' contents. Users did not experience the in-game assessment as obtrusive and uniformly perceived it as part of the game, not as part of an experiment.

Suggestions for improvement: The assigned *controls* for turning around were perceived as counter intuitive as participants had to turn the character around via one of the analog sticks of the controller in order to create similar conditions for monitor and HMD control. Another critique was the *repetitiveness* of the level during the second playthrough as participants played the same content twice, once during each presentation condition. A *flickering* of the HMD and a perceived *pixelated grid* were attributed as contributing factors for eye strain. Suggestions for further development includes the integration of an avatar or body parts such as hands or feet in order to give a frame of reference for movement tasks and to possibly increase the sense of presence further.

IV. Discussion

The general acceptance of the created environment was positive. The connection between a scientific and a game based design approach has proven useful for creating an unobtrusive in-game user state assessment by integrating it into a convincing narrative and by providing meaningful tasks. However, the evaluation yielded a number of points to be considered for improvement. First, strong levels of simulator sickness need to be addressed. Causing factors were identified as flickering of the HMD screen, stages with passive movement and non-ideal controls for turning the character around. Some of these aspects have already been improved during further development of the environment and are currently under investigation.

Second, playing the same level two times albeit during two different presentation modes proved to be repetitive. While keeping the conditions exactly constant was aimed for the scope of this study, following implementations could provide similar but not identical levels for each presentation mode. The behavioral assessment yielded no difference between conditions. Participants mentioned the competitive character of the task due to seeing their final score. Implementing further, more subtle tracking tasks or e.g. a heatmap of the players' position on x-y-z-coordinates over time are encouraged.

V. Conclusion

The introduced virtual environment was intended to be a first step in creating an approach that allows for the scientific and unobtrusive assessment of emotional reactions during interaction. A means of comparing said emotional reactions between monitor and HMD conditions was provided. Results of the conducted exploratory study suggest appropriateness of the chosen method for emotional state assessment while maintaining the users' presence within the virtual world. Users perceived the integration of a rating task via checkpoints as fitting into the narrative while the narrative and given tasks were regarded as essential for the upkeep of presence. Meaningfulness of tasks has proven to not only be a prerequisite for presence, but also a method of integrating scientific assessment tools into virtual worlds. The exploratory study yielded helpful insights for improvement and further development of the environment as well as an outlook on what to expect and what to consider for similar implementations of user state assessment during interaction in VR.

References

[1] Eber, M. (2013). How To: Skyrim and Oculus Rift with a fixed HUD.

[2] Abrash, M. (2014). What VR Could, Should, and Almost Certainly Will be Within Two Years. *Steam Dev Days*. January 15-16 2014. Seattle, WA.

[3] Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness, T.A. III, & Botella., C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour Research and Therapy, 40*(9), 983-993.

[4] Google (2015). Introducing the Expeditions Pioneer Program.

[5] Ijsselsteijn, W.A., de Kort, Y.A.W., & Poels, K. (n.d). The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games.

[6] Gibson J.J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ, Erlbaum.

[7] Lang, B. (2014, September 24). Oculus Shares 5 Key Ingredients for Presence in Virtual Reality. *Road To VR.*

[8] Despain, W. (2013). *100 Principles of Game Design*. San Francisco, CA: New Riders.

[9] Weise, M. J. (2004). *Understanding Meaningfulness in Videogames.* (Master's Thesis).

[10] Freytag, S.-C. (2016) *Untersuchung von Einflussfaktoren auf Angstreaktionen und Spielerleben beim Spiel mit und ohne headmounted VR-Displays – Design, Programmierung und Evaluation einer virtuellen Umgebung* (unpublished master's thesis). TU Berlin, Berlin, Germany.

[11] Ashforth, J. (2014). Sony: Five guidelines for effective VR design. *Interview.* Game Developers Conference Europe 2014. August 11-13, 2014, Cologne, Germany.

[12] Rothbaum, B.O., Hodges, L.F., Kooper, R., Opdyke, D., Williford, J.S., & North, M. (1995). Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *American Journal of Psychiatry, 152*, 626-628.

[13] Robillard, G., Bouchard, S., Fournier, T., & Renaud, P. (2003). Anxiety and presence during VR immersion: a comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. *Cyberpsychology, Behavior and Social Networking, 6*(5), 467-476.

[14] LaViola, J.J.Jr. (2000). A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bulletin, 32*(1), 47-59.

[15] Schandry, R. (2006*). Biologische Psychologie* (2. Aufl). Weinheim: Beltz.

[16] Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology, 3*(3), 203-220.