Evoking and Measuring Arousal in Game Settings

Jeltsje Cusveller[3], Charlotte Gerritsen[1], and Jeroen de Man[2]

[1] Netherlands Institute for the Study of Crime and Law Enforcement
De Boelelaan 1077a, 1081 HV, Amsterdam, The Netherlands
[2] Vrije Universiteit Amsterdam, Department of Artificial Intelligence
De Boelelaan 1081a, 1081 HV, Amsterdam, The Netherlands
[3] Vrije Universiteit Amsterdam, Department of Criminology
De Boelelaan 1077, 1081 HV, Amsterdam, The Netherlands
j.j.cusveller@student.vu.nl, cgerritsen@nscr.nl, j.de.man@vu.nl

Abstract. Serious games seem to be more effective if the participant feels more involved in the game. The participant should experience a high sense of presence which can be obtained by matching the level of excitement to the level of arousal a participant experiences. The level of arousal should be measured at runtime to make the game adaptive to the participant's physiological state. In this paper an experiment is presented that has as main goal to see whether it is possible to evoke arousal during different types of computer games and to monitor the physiological response. Using three online games, participants reported different levels of stress and understanding between games. Furthermore, an increase of skin conductance was found as well as a decrease in heart rate for the most difficult to understand game.

Keywords: experiment, physiological measurements, stress, serious gaming, virtual environment.

1 Introduction

Serious games are games that do not have entertainment as primary purpose. They are mainly used to for training or education purposes [8]. Examples are simulations of real-world events or processes designed for the purpose of solving a problem (e.g. Flood Sim [21], CyberCIEGE [16], Flight Gear [18]). Although serious games can be entertaining, their main purpose is to train or educate users, though it may have other purposes such as marketing or advertisement [2].

Serious games seem to be more effective if the participant feels more involved in the game. The participant should experience a high sense of presence (indicating a strong involvement of the user in the virtual environment) [13], which can be achieved by dynamically adapting the (affective) content of the game to reach a desired level of arousal [1,12]. We use the term arousal to indicate a physiological reaction to stimuli, i.e. increased heart rate and blood pressure. To obtain the 'optimal' level of arousal it is necessary to measure arousal at runtime and to make the game adaptive to it, meaning that the game should become more challenging or exciting if the level of arousal is too low and vice versa if arousal is too high [12].

S. Göbel and J. Wiemeyer (Eds.): GameDays 2014, LNCS 8395, pp. 165–174, 2014.

There already exist research projects in which arousal is measured during games [6, 11]. In this paper we build on the existing research by measuring skin conductance during three different types of games to obtain the level of arousal of the participants and investigate differences in physiological responses between games with various levels of difficulty and stressfulness. The results will be used in future research to relate changes in physiological measures to the player's mental state and develop an adaptive game for learning to deal with aggression.

In this paper, Section 3 provides some background on the research presented here and Section 3 discusses related work. In Section 4 the methods used are discussed, and in Section 5 the results are presented. Finally, Section 6 concludes the paper with a discussion.

2 Background

The research presented in this paper is part of a larger project. This project is called STRESS [14] and is joint work between the VU University Amsterdam - department of Computer Science and the Netherlands Institute for the Study of Crime and Law Enforcement.

The STRESS project [14] has as the main goal to develop an intelligent system that is able to analyse human decision making processes, and analyse the causes of incorrect decisions and inadequate emotion regulation (e.g. distress, aggression). The system will be incorporated in an ambient electronic training environment, based on Virtual Reality (VR), cf. [7, 17]. In this environment, trainees will be placed in a virtual scenario, in which they have to make difficult decisions, while negative emotions are induced. Modern Human Computer Interaction (HCI) techniques will be applied to measure aspects of their mental state (among others, stress level, emotional state, attention, and motivational state) during the scenario. This information will then be used as input for the ambient system, to determine why the trainee made certain less optimal decisions and to advise him/her how to improve this.

An important asset of the ambient approach is that the system can adapt various aspects of the training (e.g., scenarios, difficulty level, feedback) at runtime on the basis of its estimation of the trainee's mental state. In this way, both of the training goals can be fulfilled: 1) by selecting training scenarios with an appropriate context in terms of difficulty level, and providing useful feedback, the system can improve the trainee´s decision making behaviour, and 2) by selecting training scenarios with an appropriate context in terms of stress level, the system can improve the trainee´s emotion regulation skills.

3 Related Work

The work presented in this paper focuses on the physiological response in gaming. Therefore, the current section covers some related literature in the field of physiological measurements and serious gaming.

Physiological Measurements

Measuring the physiological and emotional reaction to different kinds of computer generated stimuli has been done in other research projects. In Gerritsen et al. [4] the physiological and subjective response to injustice is measured. In this work the authors performed an experiment in which the respondents had to perform a task in a virtual environment. They had to determine which of two pictures showed the more happy/sad face. If they chose correctly they moved on to the next picture combination, if they made a mistake they were set back three steps. The experiment contained two sets and in the second set the system purposely misjudged the answers. The participants were set back three combinations even if they had answered the question correctly. During the experiment the physiological responses were measured. The results showed that the level of skin conductance increased during the second set while the heart rate remained stable, indicating an increased level of arousal when feeling unjustly treated.

In the work of Brouwer et al. [3] a combination between a virtual environment and bio-neuro-feedback to help treat stress related disorders is investigated. An experiment was performed in which participants had a surveillance task in two cities. To induce stress the researchers used a bomb explosion and negative feedback. Physiological measurements were performed (heart rate variability, cortisol level and EEG) and the results indicate that the general levels of stress correlated between the participants. Besides that also associative stress was reflected in the measurements.

Serious Gaming

In the field of serious gaming the (physiological) reaction to games has been studied as well. Nacke et al. [9] present an approach to formalize evaluative methods and a roadmap for applying these mechanisms in the context of serious games. They discuss user experience (UX) and player experience models, based on which they propose a three-layer framework of GX. For each layer, they list a number of measurement methodologies and their focus is put on physiological and technical metrics for game evaluation. Their conclusion is that it depends on the kind of serious game that should be tested, which evaluative UX measure is chosen, but they do recommend affective measures for evaluating the effectiveness of motivation to play game-based learning applications or games for sports and health. While these measures might focus on emotional assessment, the authors assume that a link between positive emotion and long-term storage and recall of information in the brain exists.

In Nacke et al. [10] the authors propose a classification of direct and indirect physiological sensor input to augment traditional game control. They conducted a mixed-methods study using different sensor mappings. The results of their study show that participants have a preference for direct physiological control in games. This has two major design implications for physiologically controlled games: (1) direct physiological sensors should be mapped intuitively to reflect an action in the virtual world; (2) indirect physiological input is best used as a dramatic device in games to influence features altering the game world.

4 Methods

A computer task has been designed to measure the level of arousal in different situations while physiological measurements are registered. In the following sections, the experimental design, the participants and the hypotheses will be described.

Experimental Design

In the experiment the participants had to play three (off the shelf) computer games. These games were separated by a break. A message was presented to the participants indicating that they could continue when they felt ready. The first game (Tetris) lasted approximately four minutes, the second and third game (respectively Grid16 and Guitar Maniac) two minutes each. This of course depended on the results of the player. These games were chosen based on the anticipated pressure they would bring about. The games were presented to all of the participants in the same order.

The first game that the participants played is Tetris [20]. Tetris is a well-known puzzle game in which blocks in different shapes move from top to bottom on the screen (Figure 1). The player has to adjust the blocks so that lines will be made. A full line (from left to right) will disappear and gives the player points. If the blocks are not adjusted correctly, lines will not be made and the blocks will stack. When the blocks reach the top of the screen the game ends. Before the game started the participants received an explanation of the game and the instruction to play only the first two (easy) levels and to stop after those levels. This game was used as a baseline game. We wanted the participants to play a game which was relatively easy to do. The first two levels of Tetris are not yet very demanding and can therefore provide a baseline level in heart rate and skin conductance. The next two games were expected to be more demanding leading to a physiological response.

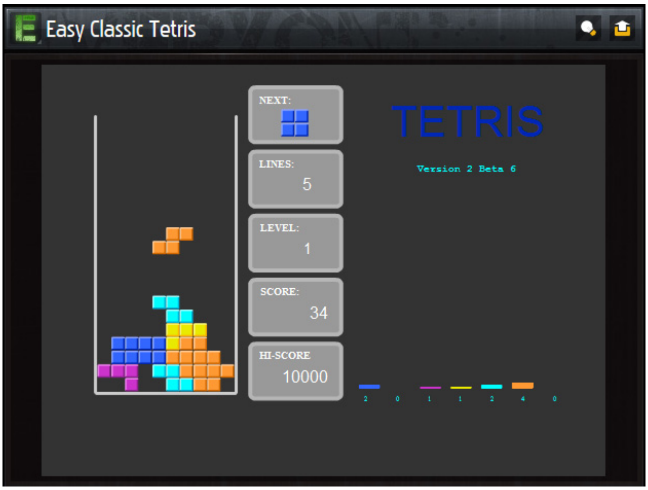

Fig. 1. Screenshot game 1: Tetris

The second game the participants played is called Grid16 [22]. In this game the participants had to play different games of skill without any explanation (see Figure 2 for a screenshot of 4 different games from Grid16). For example, in these games the participants had to avoid obstacles or had to make sure an object would not fall. The player played a game for 10 seconds and then automatically switched to another game. Meanwhile the pace increases, the games go faster and the system switches faster from one game to another game. The system continued to switch between the games until the player was *game over* in all games. The participants were instructed to play the entire game twice.

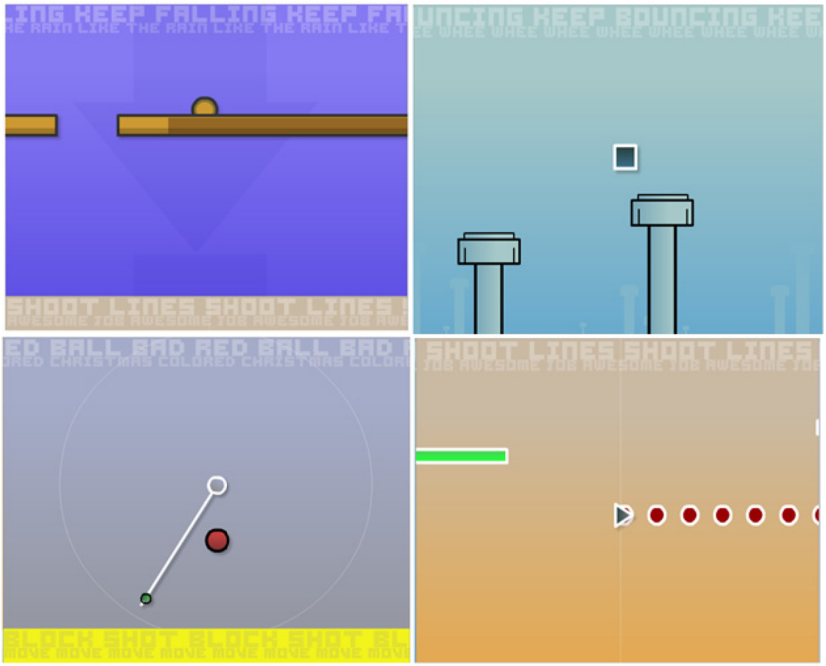

Fig. 2. Screenshots game 2: Grid16

The third game is called Guitar Maniac [19]. The player had to press the correct key on the right moment (Figure 3). On the screen the participants saw arrows move from right to left. The arrows were accompanied by numbers 1, 2, 3 or 4. When the arrows (and numbers) reached the grey part of the screen the participant had to press the key on the keyboard that corresponds to the number on the screen.

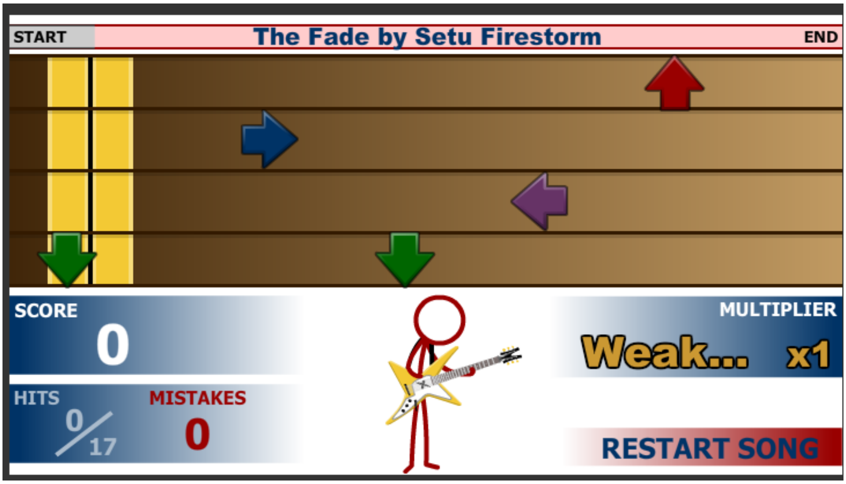

Fig. 3. Screenshot game 3: Guitar Maniac

Participants

A total of 15 participants were recruited among acquaintances of the authors. The age of these 9 women and 6 men ranged between 17 and 28 with an average age of 20.8. Each participant was fitted with the Bioplux ECG and EDA sensors measuring heart rate and skin conductance [15]. The heart rate was measured by placing three sensors on the skin near the heart. The skin conductance was measured by two sensors that were placed on the index- and middle finger. The participants were asked about their experience with games. Four participants indicated that they never play games, five participants game quite frequently. The other participants game sometimes.

Subjective Ratings

Before the experiment started the participants were asked to answer some questions to see whether the participant was relaxed before (s)he started the experiment and to find out how often the participant plays games. After each of the games the participants were asked if they understood the game, if they found the game difficult to play and if they found the game stressful to play. The answers to these questions were rated on a 7-point scale, with 0 being indicated as never/not at all and 6 as very often/very much.

Hypotheses

The expectation was that the participants would feel stress during the faster and more complicated games. It is expected that this will show in the subjective ratings of stress with a significant increase between game 1 and games 2 and 3. For Guitar Maniac, an increase in reported difficulty of the game is expected, while for Grid16 a lowered understanding is hypothesized as well. Furthermore, it is hypothesized that this increase in stress would lead to an increased level of skin conductance in game 2 and game 3. For heart rate, stating a hypothesis is more difficult, as the increase in arousal due to stress is associated with an increased heart rate. However, as stated in Section 3, similar work did not always find any changes in heart rate.

5 Results

The following sections describe the main results of the experiment. It is divided in three sections. First the results regarding the subjective questions following each game are described. Afterwards, the measurements of both skin conductance and heart rate are presented and finally correlations between the subjective and physiological results are considered.

Subjective Questions

After each game, participants answered a number of questions regarding their understanding of the game, how easy the game was to play and how stressful they experienced the game to be. Figure 4 shows for each game the average response for all 15 participants, including the standard deviation. As can be seen, Tetris was well understood and easy to play, accompanied by a low stress score for that game. Participants had more trouble understanding Grid16, found it less easy to play and also rated the game to be more stressful. Guitar Maniac was understood by the participants, but was also found difficult to play and scoring high for stress.

Statistical analysis underlines these results. A repeated one way ANOVA showed significant differences for each aspect; understanding with $F(2,28) = 56.031$, $p<0.001$, $\eta_p^2 = 0.80$, easy to play with $F(2,28) = 26.793$, $p<0.001$, $\eta_p^2 = 0.66$ and using the Greenhouse-Geisser values due to violation of the assumption of sphericity yields $F(1.44,20.22) = 26.793$, $p<0.001$, $\eta_p^2 = 0.66$ for stress. Post-hoc tests using the Bonferroni corrections shows all differences to be significant at the 0.05 level, except for the scores on easy to play and stress between Grid16 and Guitar Maniac. Further analysis has shown that there is a correlation between how easy each game was to play and how stressful participants rated the game ($r = -0.571$, $p=0.029$; $r = -0.651$, $p=0.009$; $r = -0.584$, $p=0.022$). A similar correlation was found between understanding and stress, except for guitar maniac ($r = -0.60$, $p=0.018$; $r = -0.62$, $p=0.013$; $r = -0.15$, $p=0.598$).

Physiological Measurements

Besides subjective questions, skin conductance and heart rate were measured to gain insight into the arousal of the participants. Repeated measures ANOVAs with Greenhouse-Geisser corrections showed statistical differences for both skin conductance ($F(1.266, 12.660) = 13.916$, $p=0.002$, $\eta_p^2 = 0.58$) and heart rate ($F(1.437, 15.802) = 7.818$, $p=0.008$, $\eta_p^2 = 0.42$). Post-hoc tests using a Bonferroni correction showed all differences in skin conductance to be significant at the 0.05 level, while the heart rate only significantly differs between Tetris and Grid16 as well as between Grid16 and Guitar Maniac.

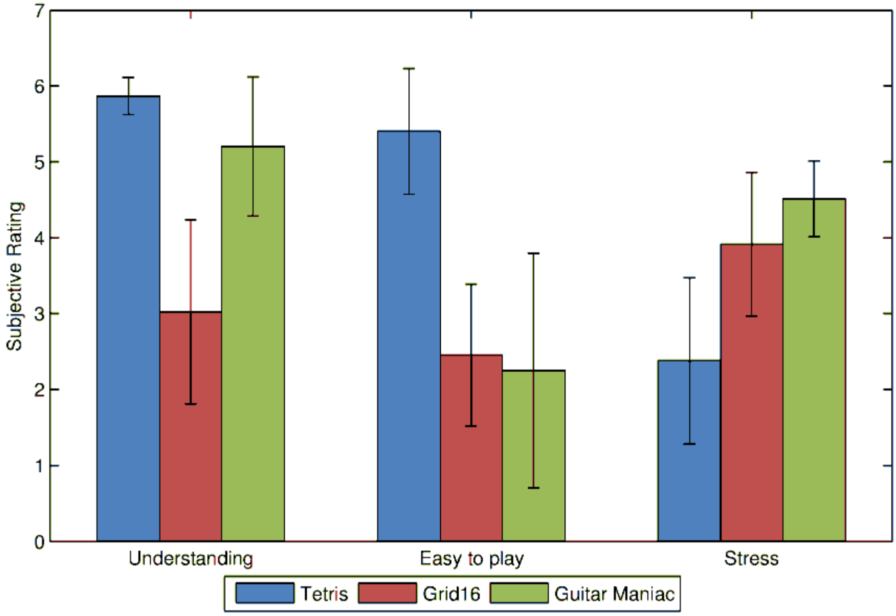

Fig. 4. Average rating (and standard deviation) to 3 different subjective questions

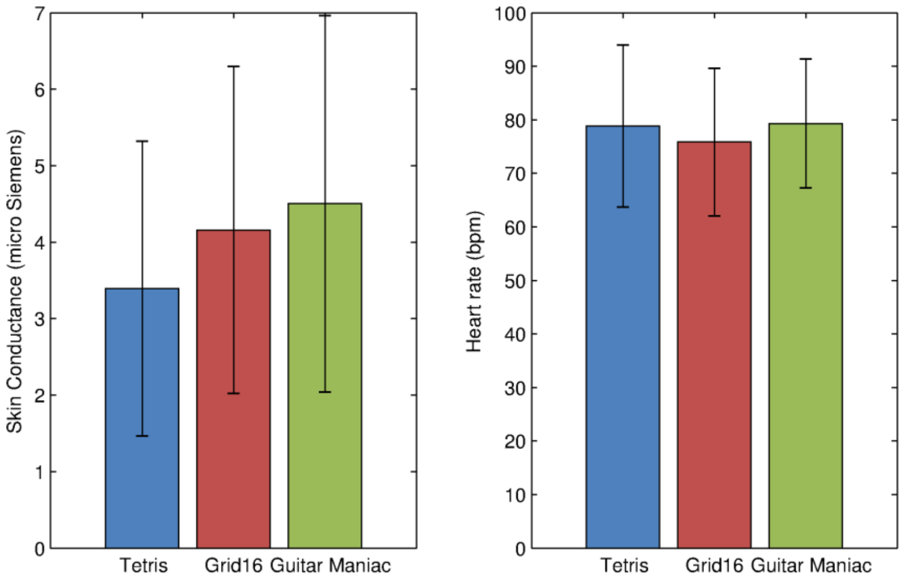

Fig. 5. Average skin conductance (left) and heart rate (right) including standard deviation for each game

Correlations between Subjective and Physiological Measurements

Finally, it has been investigated whether there were any correlations between the subjective and physiological measures. Tetris was taken as a baseline game, providing both a subjective and physiological baseline. For each of the subjective aspects and both physiological measures, the difference between Grid16 and Tetris as well as the difference between Guitar Maniac and Tetris was calculated. Spearman's rho was calculated for each combination of subjective question and physiological measurements, but resulted in no significant correlations.

6 Discussion

In the previous section, subjective and physiological results from playing three different games were discussed. Regarding the subjective questions, playing Tetris was, as expected, less stressful for the participants. Both Grid16 and Guitar Maniac were more difficult to play, making it plausible that the easier a game is to play, the less stressful it is to the participants. For Grid16, it was found that participants had trouble understanding the game, which could have caused stress. Guitar Maniac was well understood by each participant, but nonetheless hard to play and stressful. Statistically, not understanding the game did not cause a difference in the stress experienced, but on visual inspection there is a trend towards lower stress levels when this stress is (partly) caused by incomprehension.

It was hypothesized that both the skin conductance and heart rate would increase with the more stressful games. However, this was only true for the skin conductance. The heart rate dropped significantly when playing Grid16 and no differences were found with Guitar Maniac. There is some literature showing anger to be related with a decrease in heart rate [5], but it is unclear whether this applies to the situation at hand.

In the end, there was no correlation found between the answers to the subjective questions and the physiological responses. It could be that the heart rate here reflected some different aspect than stress, as it did not show the expected increase over the three games. For the skin conductance however, the increase shows a similar pattern to the increase in stress experienced by the participants, but a significant correlation between the two variables could not be found. At this point in time, it is unclear whether the skin conductance is indeed not directly correlated with the stress experienced, or that a weaker correlation between the two variables is obscured by a low number of participants.

Another aspect not considered in this work yet is personality traits. Each participant undertook a personality tests and initial analyses show some correlations between personality traits and the participants' physiological responses. A more in-depth analysis of these personality traits might reveal underlying correlations between stress and physiological responses.

Acknowledgement. The authors are grateful to Tibor Bosse for his feedback and fruitful discussions. This research was supported by funding from the National Initiative Brain and Cognition, coordinated by the Netherlands Organisation for Scientific Research (NWO), under grant agreement No. 056-25-013.

References

1. Baños, R.M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., Rey, B.: Immersion and Emotion: Their Impact on the Sense of Presence. CyberPsychology & Behavior 7(6), 734–741 (2004)
2. Bergeron, B.P.: Developing Serious Games. Charles River Media, Massachusetts (2006)
3. Brouwer, A.M., Neerincx, M.A., Kallen, V.L., van der Leer, L., ten Brinke, M.: EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. Journal of Cybertherapy & Rehabilitation 4(1), 21–34 (2011)
4. Gerritsen, C., de Man, J., van der Meij, J.: Physiological and subjective response to injustice: the effects of unjust evaluations on physiological responses and subjective experiences. In: Proceedings of the Seventh International Workshop on Human Aspects in Ambient Intelligence, HAI 2013. IEEE Computer Society Press (2013) (to appear)
5. Hassett, J.: A primer of psychophysiology. WH Freeman, San Francisco (1978)
6. Kotsia, I., Patras, I., Fotopoulos, S.: Affective gaming: beyond using sensors. In: Proceedings of the Fifth International Symposium on Communications, Control and Signal Processing, ISCCSP 2012 (2012)
7. Marsella, S., Gratch, J.: Modeling the influence of emotion on belief for virtual training simulations. In: Proceedings of the 11th Conference on Computer-Generated Forces and Behavior Representation, Orlando, FL (May 2002)
8. Michael, D.R., Chen, S.L.: Serious Games: Games That Educate, Train, and Inform, Muska & Lipman/Premier-Trade (2005)
9. Nacke, L.E., Drachen, A., Goebel, S.: Methods for Evaluating Gameplay Experience in a Serious Gaming Context. International Journal of Computer Science in Sport 9, 2 (2010)
10. Nacke, L.E., Kalyn, M., Lough, C., Mandryk, R.L.: Biofeedback Game Design: Using Direct and Indirect Physiological Control to Enhance Game Interaction. In: Proc. CHI 2011 (2011)
11. Nacke, L.E., Mandryk, R.L.: Designing affective games with physiological input. Proceedings of Fun and Games (2011)
12. Nogueira, P.A., Rodrigues, R., Oliveira, E., Nacke, L.E.: Guided Emotional State Regulation: Understanding and Shaping Players' Affective Experiences in Digital Games. In: Proceedings of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, pp. 51–57. AAAI (2013)
13. Steuer, J.: Defining virtual reality: Dimensions determining telepresence. In: Biocca, F., Levy, M.R. (eds.) Communication in the Age of Virtual Reality, pp. 33–56. Lawrence Erlbaum Associates, Hillsdale (1995)
14. `http://stress.few.vu.nl/`
15. `http://www.biosignalsplux.com/`
16. `http://www.cisr.us/cyberciege/`
17. `http://www.ehow.com/list_5982829_benefits-police-trainingsimulators_.html`
18. `http://www.flightgear.org/`
19. `http://www.kongregate.com/games/shinki/super-crazy-guitar-maniac-deluxe-3`
20. `http://www.newgrounds.com/portal/view/235136`
21. `http://playgen.com/play/floodsim/`
22. `http://www.spelletjes.nl/spel/Grid-16.html`