Assessing Game Experience: Heart Rate Variability, In-Game Behavior and Self-Report Measures

Conference Paper · December 2014

DOI: 10.1109/QoMEX.2014.6982334

3 authors:

Elena Patricia Nunez Castellar

Ghent University

Kimmo Oksanen

University of Jyväskylä

Jan Van Looy

Ghent University

ASSESSING GAME EXPERIENCE: HEART RATE VARIABILITY, IN-GAME BEHAVIOR AND SELF-REPORT MEASURES

Elena Nunez Castellar[a], *Kimmo Oksanen*[b], *Jan Van Looy*[a].

[a] Department of Communication Studies, iMinds-MICT, Ghent University, Belgium
[b] Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland

ABSTRACT

Assessing game experience by means of recordings of physiological reactions elicited during game play is a technique that has gained popularity in recent years in the field of digital games research. However, since physiological signals are typically linked to several psychological processes, the use of some measures such as cardiac activity or heart rate (HR) remains problematic. The goal of the present study is to investigate to what extent game logs and self-report measures of game experience have a predictive value for heart rate variability during game play. Our results showed that the accurate registration of in-game behaviors by means of game logs carries the potential of providing richer information for the interpretation of variations in cardiac activity than only post-hoc self-report measures.

Index Terms— heart rate, game logs, self-report measures, game experience.

1. INTRODUCTION

Over the past few years, interest in using psychophysiological techniques to investigate quality of experience (QoE) during media exposure has grown considerably. One of its potential advantages compared to other techniques such as self-report, apart from being an objective measure, is the fact that it has a high temporal resolution (e.g in the millisecond range). However, although psychophysiology has previously been used in game experience research [1], a number of questions remain regarding the use of some measures such as cardiac activity and heart rate (HR).

Ekman and colleagues [2], for example, point to difficulties in interpreting HR due to the central role of blood circulation in the human body. The heart is regulated by many processes and systems and therefore it is often difficult to link HR signals to a single psychological process. In fact, HR signals have been used as an index of a broad spectrum of psychological processes such as attention, cognitive effort, stress, arousal, valence, and orientation reflex in the field of media psychology [3]. In digital games research, changes in heart rates are frequently observed when recorded, but the source of the HR signal's variations remains largely unknown.

Hence, self-report measures of gaming QoE are frequently used to support the interpretation of HR signals. In fact, it is commonly accepted that by combining these two indexes, QoE can be assessed more accurately [2]. However, self-report measures have limitations which are often neglected. For instance, the ability to provide accurate self-reports of thoughts and emotions relies on psychological processes that can differ among subjects such as memory, attention, and introspective ability.

Recently, the use of game logs, has started to gain popularity in the field of media research [4]. Objective indexes of game performance, or in-game behavior, combined with self-report are also thought to carry the potential of providing a more accurate measure of gaming QoE. Its ability to complement and support the interpretation of psychophysiological data is something that has hitherto been largely overlooked in digital games research however [4].

The present study looks into the ways in which different types of QoE measurement in games research can reinforce and cross-validate one another. Particularly we explored how log analysis can be used to support the interpretation of heart rate data. Thereby we first investigate whether self-report scores of game experience and in-game behaviors obtained by means of automatic data logging, have a predictive value for heart rate variability (HRV) changes during game play. Second, since the game was played in groups, we also investigate whether group actions during the game modulates the averaged HRVs and the self-report game experience scores. To this end, participants played the multiplayer collaborative game Game Bridge while HRV and game logs were simultaneously recorded.

2. METHOD

2.1. Participants

Here we draw on data of 55 vocational students and 10 teachers (40 male and 15 female), which were assigned to 11 gaming groups with five participants. Due to missing data, one participant had to be excluded. The data reported here include 54 subjects (N=54). The results game experience questionnaire (GEQ) and HRV grand averages from this sample have already been reported elsewhere [13]. However, this report did not examine in-game behaviors based on the game logs and its relationship with HRVs. Also no statistical analyses were conducted for HRVs (only descriptive data was reported).

2.2. Collaborative game

Game Bridge is a multiplayer collaborative game, in which maximum 5 players are connected to each other via a game server that runs in a virtual world in which game actions take place (see Figure 1). The goal of the game is to emphasize the importance of shared work between professionals of different areas and relevance in real life, to create productive collaborative knowledge construction among the players and to expand their awareness of human sustainability.

Figure 1. Game Bridge screenshot

The game's story is that players work as volunteers at a charity concert for human sustainability. Through three pedagogically scripted phases (or levels), they are supposed to make sure that customers are satisfied and that everything is ready for a band's gig. The first phase (Gate) demands accurate coordination between the players, the second (Restaurant) the accurate combination of players' expertise and the final phase (Stage) requires solving cognitive conflict. Since the game has been designed to promote collaborative activities, each phase requires players to complete a multiplayer puzzle that demands commitment for its successful completion (for a more detailed description see [5,13]).

2.3. Procedure and Measures

The study was conducted in a classroom setting. The three phases were completed in about two hours and a half while logs of in-game behavior and HRV (heart rate variability) were simultaneously recorded. HR data were recorded using Firstbeat Bodyguard (Firstbeat Technologies OyTM, Jyväskylä, Finland) equipment by attaching the device directly to the skin (two points) and with a sampling rate of 1000 Hz. Before starting, participants were requested to sit back and relax during eight minutes to record a baseline period for the heart rate measures. At the end of the experiment, each individual filled out the Game Experience Questionnaire (GEQ), which was developed, implemented and validated during the "Fun of Gaming" (FUGA) project. Complete information about the GEQ can be found in [6], but relevant for the present study is the core module of the GEQ.

2.4. Data analysis

2.4.1 Heart rate

As a parameter for the analyses of the heart rate data, the root mean square of successive differences in RR intervals was used (RMSSD, see [7]). Briefly, RMSSD is time-domain tool used to assess heart rate variability that computes the successive differences being neighboring RR intervals. Specifically, it calculates the square root of the mean of the squared differences between successive RR intervals over time. The HRV data was processed with the software "Firstbeat HEALTH" (Firstbeat Technologies OyTM, Jyväskylä, Finland), artifacts were removed and the RMSSD was calculated with the software's formula for minute time spans. Afterwards, the baseline was calculated per participant using the last 4 minutes of the resting phase and this value was used to get baseline-corrected HRV signals. Finally, the signals were divided into three parts according to the three phases of the game.

2.4.2 Logs: in-game data

The game is divided in three phases and accordingly different actions should be performed. In the first phase (Gate) players need to open a gate to the festival area by entering a password to the electronic lock. Every player has gotten their own part of the password. The fragments of the password must be entered in the correct order. During the second phase (Restaurant) the players are requested to keep customers and band members satisfied by serving them in the catering area. Players should integrate and synchronize their individual tasks in a timely manner (complementary actions) to keep the customers satisfied. Finally, during the third Phase (Stage) players are requested to share their individual information with each other in order to get a shared understanding of all the available information. Specifically they should identify each band member and organize the band's equipment in the right place on the stage. Identification was based on the tips provided by roadies and pictures on the boxes.

Overall we identified four types of actions in the log data for the three phases:

- Actions performed: tasks performed without or with a specific task given. For example having a forced break, picking up a piece of garbage or taking the order from the customer.
- Failed: given task failed, for example because missing the time limit.
- Succeed: given task succeeded for the player, for example taking an order from the customer.
- Task given: tasks performed with a specific task given

Total counts for each of these types of actions were used to conduct the statistical analyses.

2.4.3 GEQ scores

GEQ core module scores were computed for each of the 7 different dimensions of game experience: flow, sensory and imaginative immersion, competence, challenge, positive and negative affect, and tension. Example statements of the dimensions are as follows, “I felt completely absorbed” (flow), “It was aesthetically pleasing” (immersion), “I felt skillful” (competence), “I had to put a lot of effort into it” (challenge), “I enjoyed it” (positive affect), “ I felt bored” (negative affect), “I felt pressured” (tension). GEQ was collected at the end of the experiment, and not after each of the different stages, in order to prevent that the game flow breaks. Dimension scores were computed as the average value of its likert-type items, which are scored in a range from 1 (not at all) to 5 (extremely). The other modules have not been analyzed in the present study given that previous research has shown that in general they are highly interrelated with the core module [8].

2.4.4 Individual vs. Group analysis

We first conducted correlation analyses between heart rate variability, in-game behaviors and self-report measures of game experience for each of the three phases of the game. For this purpose, the individual scores (each participant was considered a data point) were included in the analyses.

Afterwards we explored group differences. We conducted a one-way analysis of variance (ANOVA) to explore differences between groups. Next we conducted correlation analyses but on the group scores (each group was considered as a data point) for the three indexes of interest. For all the tests, a significance level of 0.05 was used.

3. RESULTS

3.1. Individual analysis

Descriptive statistics for each of the three measured quantities (GEQ, HR, logs) were conducted. The average GEQ scores and standard deviations for each of the dimensions were calculated; flow (M = 2.9, SD = 0.9), sensory and imaginative immersion(M = 2.8, SD = 0.8), competence (M = 3.1, SD = 0.6), challenge (M = 2.3, SD = 0.6), positive (M = 3.4, SD = 0.8) and negative affect (M = 1.9, SD = 0.7), and tension (M = 2.0, SD = 0.7). The average HR and standard deviations for each phase were also calculated; first phase (M = 10.9, SD = 12.4), second phase (M = 10.5, SD = 12.1) and third phase (M = 6.3, SD = 10.1). The average actions and standard deviations per category were also calculated Actions performed (M = 178, SD = 120.3), Succeed (M = 50, SD = 21.8), Failed (M = 2, SD = 1.8) and Tasks given (M = 55, SD = 22.5).

Afterwards correlation analyses were conducted. The results of correlation analyses between HRV and in-game logs indicate a statistically significant association between the number of actions performed during the game and HRVs during the third phase ($r(54) = -.32$, $p < .05$). Also a close to significant association during the second phase of the game ($r(54) = -.26$, $p = .05$) was revealed. The negative correlation indicates HRVs decreases when more actions are performed in the game. Importantly only the number of actions, and not the sort of outcomes (such as successfully performed actions and failures), were associated with changes in HRVs.

Besides the relationship with in-game logs, we also explored whether HRV changes were associated with scores of the game experience dimensions included in the GEQ. The results showed that only the self-report scores for the variable “tension” were significantly associated with the HRV changes during the second phase of the game ($r(54) = -.27$, $p < .05$) and marginally significantly associated during the first and third phase ($r(54) = -.23$, $p < .10$ & $r(54) = -.25$, $p < .10$). The negative correlation indicates that a larger decrease in HRVs was associated with higher self-report scores of feeling “tension” during the game (see table 1).

	HRV			GEQ scores						
	phase 1	phase 2	phase 3	competence	flow	tension	challenge	neg affect	pos affect	inmersion
HRV										
phase 1	1	,872**	,754**	,128	,098	-,232	-,104	-,088	,166	-,114
phase 2		1	,919**	,120	,069	-,270*	-,118	-,105	,177	-,148
phase 3			1	,132	,058	-,251	-,067	-,062	,152	-,156
GEQ scores										
competence				1	,502**	-,153	,061	-,360**	,665**	,584**
flow					1	,054	,533**	-,314*	,646**	,719**
tension						1	,452**	,678**	-,310*	-,204
challenge							1	,172	,256	,397**
neg affect								1	-,596**	-,571**
pos affect									1	,807**
inmersion										1

Note. **. Correlation is significant at the 0.01 level *Correlation is significant at the 0.05 level

Table 1. HRV and self-report scores of game experience

3.2. Group analysis

3.2.1 Group HRVs and in-game behavior

HRVs were subjected to a one-way ANOVA. The main effect of Group was significant ($F(10,54) = 101,43$, $p < .001$), suggesting that HRV averages, significantly differed between groups (blue line Figure 3).

To assess whether HRV changes per group were associated with the type and number of behaviors executed during the game, we conducted regression analyses. The results can be observed in Figure 2. Our results revealed that the unique in-game behavior that shows a significant correlation with HRV changes is the number of actions performed by the group ($r(54) = .59$, $p < .10$). This indicates that the larger the number of actions performed by the group, the larger the HRV deceleration (see Figure 3, negative is plotted upwards for HRVs).

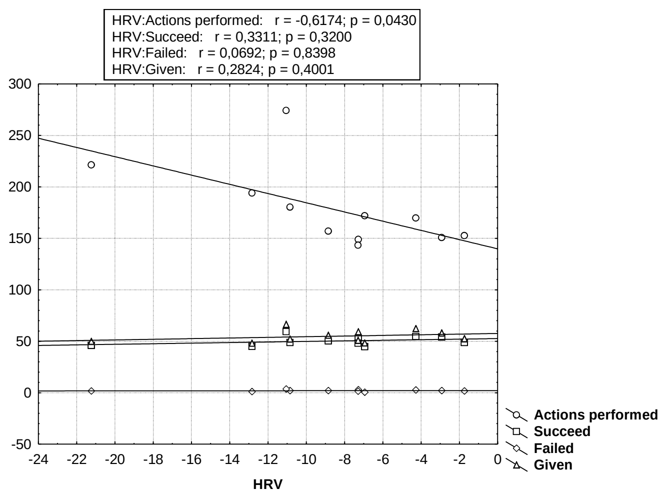

Figure 2. Correlation between HRVs and in-game behavior based on game logs

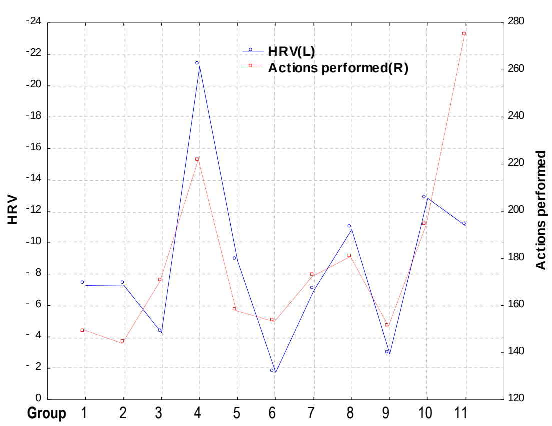

Figure 3. Double Y graph: averaged HRVs (left) and the number of actions performed (right) during the game per group (x axis).

4.2.1 Group HRVs and its association with self-report measures of game experience

We also conducted regression analyses to investigate whether HRV changes per groups were associated with the averaged self-report scores of game experience. The results did not reveal any significant or marginally significant association (all p values > .1)

5. DISCUSSION

Although the present study was explorative in nature, our results point to three interesting findings.

First, the correlation analyses between HRVs and in-game behavior revealed that, when considering individual data, a significant association can be found between the number of actions performed in the game and HRV changes. Importantly, the outcome of these actions (failures vs. successful actions) did not yield significant results. This result might indicate that HRVs deceleration was driven by general game demands rather than by emotional arousal (emotions elicited by positive vs. negative game outcomes). This finding is inconsistent with previous research suggesting that HR covaries primarily with emotional arousal [9] but is in line with the hypothesis that HRVs are associated with mental workload during game play [10] , and findings showing that HRV decreases with high mental effort [11].

Secondly, in line with the observation that HRVs seem to be a general index of game demands, our results revealed that out of the seven investigated dimensions of game experience, the self-report scores for the dimension *Tension* were the only score significantly associated with HR deceleration. As this dimension included items about how tense, how frustrated or pressured the participants felt during the game, further studies could investigate whether this dimension maps for negative affect [12] or for general demands. Given that we found HRVs were not associated with objective measures of negative outcomes (failures), we believe that the correlation with the items of the "*Tension"* dimension of the GEQ is primarily driven by the mental workload.

A third relevant finding of the present study is that, when analyzing averages by groups of the three indexes, there was no association between the HRVs and the averaged self-report scores of game experience per group (see figure 1); yet there was a significant correlation between the HRVs and the number of actions performed as registered in the game logs (see Figure 1b). These results show that game logs can be of great importance as side data to interpret HRVs changes.

Taken together, our results show that game logs are complementary and have the potential to provide rich information for the interpretation of physiological data. Since there is a lack of a commonly accepted theory on how digital game experience arises [1], researchers are advised to consider not only self-reported measures of game experience but also in-game behavior as an important source of information. Finally another important advantage of using game logs is that it is a non-intrusive measure that imposes no burden on the user, which can be an important aspect when assessing QoE.

6. REFERENCES

[1] J. M. Kivikangas, I. Ekman , G. Chanel, S. Järvelä, M. Salminen, B. Cowley, ... & N. Ravaja, Review on psychophysiological methods in game research. *Proc. of 1st Nordic DiGRA*, 2010.

[2] I. Ekman, G. Chanel, S. Järvelä, J.M. Kivikangas, M. Salminen, & N. Ravaja, Social Interaction in Games Measuring Physiological Linkage and Social Presence. *Simulation & Gaming*, *43*(3), 321-338, 2012.

[3] N. Ravaja, Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology*, *6*(2), 193-235, 2004.

[4] G. Wallner, & S. Kriglstein, Visualization-based analysis of gameplay data: A review of the literature. *Entertainment Computing*, 4(3), 143-155, 2013.

[5] K. Oksanen, R. Hämäläinen, B. Mannila,, & T. Manninen, Designing and investigating game tasks for supporting collaborative learning. In B. Meyer (Ed.), Proceeding of the 4th European Conference on Games-Based Learning (pp. 291–298). UK: Academic Publishing, 2010.

[6] W. IJsselsteijn, K. Poels, Y.A.W. de Kort, The Game Experience Questionnaire: Development of A Self-report Measure to Assess Player Experiences of Digital Games, TU Eindhoven, Eindhoven, 2008.

[7] J.J. Allen., A.S. Chambers, & D. N. Towers, D. N. The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological psychology*, *74*(2), 243-262, 2007.

[8] B.J. Gajadhar, H.H. Nap, Y.A. de Kort, & W.A. IJsselsteijn, Out of sight, out of mind: co-player effects on seniors' player experience. In *Proceedings of the 3rd International Conference on Fun and Games*, pp. 74-83, 2010.

[9] N. Ravaja, T. Saari, M. Turpeinen, J. Laarni, M. Salminen, & M. Kivikangas. (2006). Spatial presence and emotions during video game playing: Does it matter with whom you play?. *Presence: Teleoperators and Virtual Environments*, *15*(4), 381-392, 2006.

[10] B. Cowley, N. Ravaja, & T. Heikura. Cardiovascular physiology predicts learning effects in a serious game activity. *Computers & Education*,*60*(1), 299-309, 2013.

[11] D. W., Rowe, J. Sibert & D. Irwin.. Heart rate variability: Indicator of user state as an aid to human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 480-487). ACM Press/Addison-Wesley Publishing Co, 1998.

[12] IJsselsteijn, W., Poels, K., & de Kort, Y. A. W. The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games. Eindhoven: TU Eindhoven. *FUGA Deliverable* D3.3. Technical Report, 2008.

[13] K. Oksanen & R. Hämäläinen. Assessing Game Experiences Caused by Educational Collaborative Game. In Proceeding of the 5th European Conference on Games-Based Learning (pp. 431–437). UK: Academic Publishing, 2011.