Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

Klarkowski, Madison, Johnson, Daniel, Wyeth, Peta, Phillips, Cody, & Smith, Simon
(2016)
Psychophysiology of challenge in play: EDA and self-reported arousal. In *CHI 2016 Workshop: HCI and Autonomous Vehicles: Contextual Experience Informs Design*, 7-8 May 2016, Stanford, Mountain View, and San Jose.

This file was downloaded from: https://eprints.qut.edu.au/105321/

© 2016 The Author(s)

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

https://doi.org/10.1145/2851581.2892485

Psychophysiology of Challenge in Play: EDA and Self-Reported Arousal

Madison Klarkowski
Queensland University of Technology
Brisbane, Australia
m15.clark@qut.edu.au

Daniel Johnson
Queensland University of Technology
Brisbane, Australia
dm.johnson@qut.edu.au

Peta Wyeth
Queensland University of Technology
Brisbane, Australia
peta.wyeth@qut.edu.au

Cody Phillips
Queensland University of Technology
Brisbane, Australia
c1.phillips@qut.edu.au

Simon Smith
Queensland University of Technology
Brisbane, Australia
simon.smith@qut.edu.au

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).
CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892485

Abstract
Measuring the video game player experience is a distinctly challenging task. As the experience of 'fun' in games is imprecise and multi-faceted, various psychological and experiential phenomena have been investigated in an effort to evaluate and quantify aspects of the player experience. Psychophysiology provides a useful lens through which to objectively and quantitatively measure and evaluate these phenomena. This study reports current electrodermal activity (EDA) findings from a large-scale ongoing study investigating the psychophysiology of play using electrodermal activity, electroencephalography, electromyography, and electrocardiography. Initial EDA results point to greater arousal the more challenging the play experience. Findings also indicate that EDA potentially reports arousal with greater real-time accuracy than a subjective arousal measure. Ultimately, with this work, we aim contribute to a greater understanding of the psychophysiological evaluation and impact of play.

Author Keywords
Psychophysiology; physiology; EDA; biometrics; video games; games; play experience.

ACM Classification Keywords
K.8.0 Personal Computing: General- games.

Introduction

Psychophysiology is an objective-quantitative method that allows for the investigation and evaluation of the physiological impact of the play experience (PX). To do so, psychophysiology obtains physiological signals, recorded by biometric devices, as a method of measuring a user's mental and emotional state [1]. These signals are the physiological response to a person's psychological state; perhaps the most conspicuous example of this is increased heart-rate in response to fear, excitement, or general arousal [1]. The use of these devices allows for a real-time measurement of the PX, and can be correlated with subjective methods to determine accuracy or aid in the interpretation of results [5]. Additionally, as psychophysiology measures involuntary physiological response, the measurement is free from some of the limitations associated with other measures - for example, interruption of gameplay for interview purposes, or reliance on participant recall and interpretation of survey items [5]. Although there hasn't been a unified approach to psychophysiological analysis of the PX, it has been recommended as an analysis strategy to concentrate on [2].

Games literature has explored the psychophysiological impact of flow [4] [9], avatar choice [7], immersion [9], violent content and difficulty [6], social play [8], sonic user experience [10], mood [13], and reward [11], amongst other concepts. Generally, and through necessity engendered by study scope, psychophysiological research explores a single psychological construct at a time, using one or two physiological measures. Additionally, psychophysiological studies in games are often limited by small sample sizes [5]. In a review of psychophysiological methods in games research, Kivikangas et al. states, *"Thus, we have a number of separate results for many separate research questions, but very little accumulated knowledge that could be used for answering more precise research questions or for creating theoretical syntheses."* [5]

To address this gap, we have designed a large-scale psychophysiological study of prominent experiential phenomena. This study employs electroencephalography, electrocardiography, electromyography and electrodermal activity to evaluate the experience of play through self-reported affect, autonomy, competence, valence, arousal, dominance, flow, presence, boredom, and enjoyment. To address the recurring obstacle of small sample sizes, this study will collect data from 120 participants. This study represents a step towards creating a unified approach to psychophysiological PX evaluation, with a focus on establishing an accessible and robust psychophysiological methodology for PX research and playtesting.

This paper reports initial electrodermal activity (EDA) and self-report arousal findings from the first sixty-one datasets collected in the study. Both the psychophysiological and the subjective response are evaluated across three game conditions that evoke an optimal gameplay experience (medium challenge) and sub-optimal gameplay experiences (too easy, boring; too hard, overwhelming). Familiarity with psychophysiological markers associated with both successful and unsuccessful gameplay designs would represent a contribution to understanding the PX in research, and would assist in providing a useful tool for video game developers and designers.

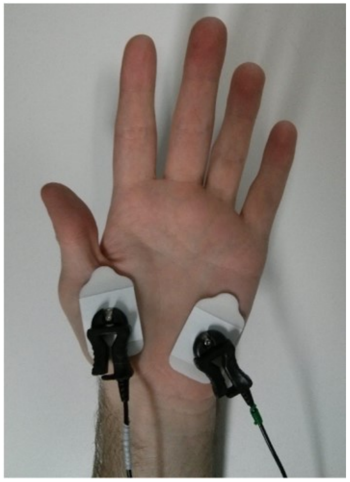

Figure 1: EDA electrodes applied to the palm.

Electrodermal Activity (EDA)

EDA (see Figure 1) is the measurement of eccrine sweat gland activity in the palms of the hands and soles of the feet [14]. The eccrine sweat glands are unique in that, unlike other sweat glands that mainly moderate body temperature, eccrine glands respond primarily to cognitive and emotional stimulation [14]. EDA measures the arousal to this stimulation through cutaneous electrodes, with heightened EDA indicative of greater experienced arousal [1]. As such, EDA is a measure of physiological arousal – the intensity with which an emotion is experienced.

In Games Research

As measures of EDA are generally easily deployable, EDA represents one of the most popular psychophysiological measures in PX research. Greater EDA response has been found in participants during a flow-inducing play experience than in immersive and boring play experiences [9] and when playing competitive video games with friends than with strangers [8]. Increased EDA response was found when a player killed or wounded enemy opponents, or their own player-character was killed or wounded [12]. Interestingly, Kneer, Elson & Knapp found no effect on psychophysiological arousal (as measured by EDA) as a consequence of difficulty or violent content [6].

Study Description

This study represents partial initial findings from a greater program of research investigating the psychophysiology of play. The current paper investigates electrodermal response to three different game conditions. These game conditions were developed within *Left 4 Dead 2*, a first-person zombie shooter. Each of these conditions were designed with the intent to thwart or promote an enjoyable game experience through the manipulation of challenge; this was to help identify the psychophysiological experience of a 'successful' gameplay experience versus an 'unsuccessful' (that is, too easy or too difficult) one. In between each play session, participants would answer digitised questionnaires about their experience of autonomy, competence, presence, arousal, valence, autonomy, dominance, affect, enjoyment, flow, fun, boredom, and perceived challenge-skill balance. Arousal is reported in this paper. The psychophysiological measures taken were EEG, EDA, EMG, and ECG; reported in this study is EDA. The study took a tonic approach, investigating averaged EDA across each game condition.

Participants

Sixty-six participants (fifty male), aged 17 – 38 (mean age of 23.28, *SD* = 4.69), volunteered for the study. On a Likert scale of 1 – 7, with '7' representing 'extremely experienced' and '1' 'not at all experienced', participants self-rated as an average of 5.93 (*SD* = 1.35) for 'general experience with video games' and 5.18 (*SD* = 1.81) for 'experience with first-person shooters'.

Measures

EDA

EDA was measured using BIOPAC EL507 snap-on electrodes pre-filled with isotonic gel. The electrodes were attached in a typical bipolar placement on the thenar and hypothenar muscle sites (see Figure 1), which feature a high concentration of eccrine sweat glands [14]. Participants washed their hands with hypoallergenic soap prior to electrode attachment. The electrodes were secured with medical tape. A grounding

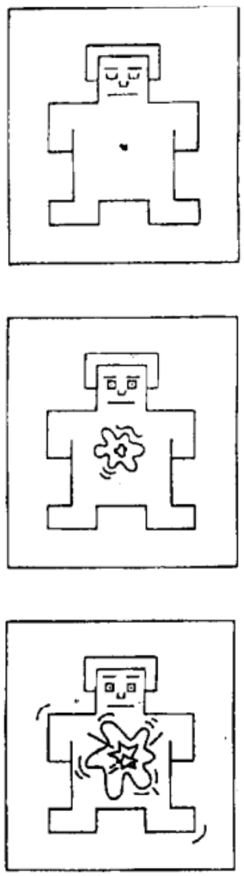

Figure 1: Example figures from the SAM Arousal scale. Top is least arousal, middle is moderate arousal, and bottom is most arousal. [3]

electrode was placed on the participant's forehead alongside facial EMG electrodes, for which the results are not reported in this paper.

Self-Assessment Manikin

The Self-Assessment Manikin (SAM) is a pictorial self-report measure of valence, arousal, and dominance [3], in which participants select a figure that most closely represents their current emotional state. Participants self-rated on all three scales after each play session, although only arousal is reported in this paper. Arousal is represented by a growing 'explosion' within five figures, in which low arousal is represented by a small dot and closed eyes, and high arousal is represented with a large, shaking explosion and raised eyebrows. See Figure 2 for example figures.

Game

As this research aims to position psychophysiological measures as applicable and beneficial for contemporary playtesting environments, and reflective of contemporary play experiences, it was essential to employ a video game typical of a popular modern title. The game chosen for this study was Valve Corporation's *Left 4 Dead 2*.

Three play conditions that corresponded with low challenge ('Boredom'), medium challenge ('Balance'), and high challenge ('Overload') created within *Left 4 Dead 2.* In all conditions, players were required to pick up gas canisters throughout the level. The canisters were marked clearly in the game world so that finding them was not challenging.

In the Boredom condition, this task was carried out with no enemy resistance. The Balance condition featured standard enemy agents that dynamically matched the player's in-game performance, ensuring the condition would not be too easy or too hard for the player. Finally, the Overload condition featured continually spawning enemies that hit for ten times the damage of the Balance condition, low player health, and limited ammunition reserves, impeding completion of the level.

Process

Each experiment session took approximately 120 minutes, including a forty-five minute setup period for the psychophysiological instruments. Participants would play a custom tutorial for the game so as to familiarise themselves with the controls and mechanics. Participants would then play three ten-minute gameplay session, answering questionnaires about their play experience between each session.

A repeated measures experiment design was employed to reduce learning effect. Consequently, thirty-three participants played the Balance condition first, and thirty-three participants played the Boredom condition first. The Overload condition was played last so as to prevent participant frustration influencing play experience in the Balance and Boredom conditions.

A two-minute baseline at the start and end of the experiment, and in between each play session, was also implemented. The baselines allowed for participants' physiological response to 'reset' prior to each play condition, reducing the possibility of residual physiological effect from experimenter interaction or surveys.

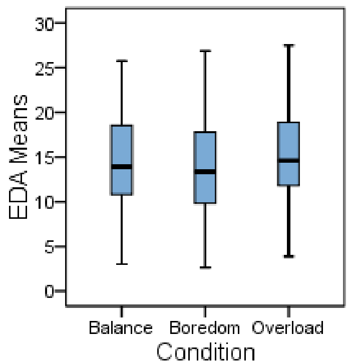

Figure 3: EDA physiological arousal means for each play condition.

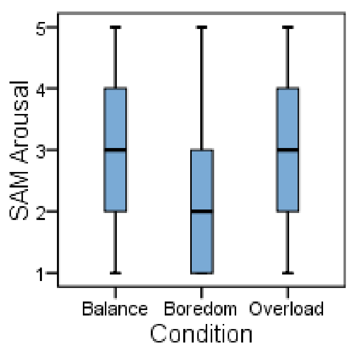

Figure 4: Self-reported SAM arousal means for each play condition.

Data Treatment

The EDA data was analysed using 10-second epochs through epoch analysis performed in Biopac's *AcqKnowledge* data acquisition and analysis software. Ten minutes, or sixty epochs, of data was acquired from each play session, representing thirty minutes of EDA data per participant.

Data Loss

All data was visually scanned for movement artefacts, with any contaminated data removed from the dataset prior to statistical analysis. Five datasets were identified as entirely compromised and removed from the final sample. This was likely the result of loss of electrode contact. As such, statistical analysis was performed on sixty-one datasets.

Results

EDA

A repeated-measures ANOVA was conducted on the EDA data, using the experimental condition as the within-subjects factor. There were no outliers in the data, as assessed by inspection of a boxplot. Assumptions of normality were subsequently satisfied in all instances, as assessed by the Shapiro-Wilk's test ($p > .05$). Mauchly's test of sphericity indicated that the assumption of sphericity had been violated ($W = .836$, $\chi^2(2) = 10.421$, $p = .005$), and so a Greenhouse-Geisser adjustment ($\varepsilon = .859$) was applied.

A significant main effect of the experimental condition on EDA was revealed ($F(1.718, 101.334) = 16.951$, $p < .001$, $\eta p2 = .223$). Post-hoc analysis with a Bonferroni adjustment revealed that participants produced significantly less EDA response in the Boredom condition ($M = 13.84$, $SD = 5.90$) than in both the Balance ($M = 14.58$, $SD = 5.52$, $p = .001$) and Overload ($M = 15.20$, $SD = 5.68$, $p < .001$) conditions, and significantly less EDA response in the Balance Condition than the Overload Condition ($p = .019$). Please refer to Figure 3.

SAM

A repeated-measures ANOVA was conducted to determine whether there were statistically significant differences in self-report SAM Arousal between three experimental conditions (Balance, Boredom, Overload). There were no outliers in the data, as assessed by inspection of a boxplot. Normality checks revealed evidence of moderate positive skew on SAM arousal. All analyses were run with transformed and non-transformed versions of the data. No differences in patterns of results emerged, hence non-transformed results are reported here for ease of interpretation. The assumption of sphericity was not violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 3.789$, $p = .150$).

A significant main effect of the experimental condition on SAM Arousal was revealed, $F(2,128) = 43.530$, $p < .001$. Post-hoc analysis with a Bonferroni adjustment revealed that participants reported significantly higher arousal in the Balance condition ($M = 3.02$, $SD = 1.125$) than the Boredom condition ($M = 2.03$, $SD = 1.045$, $p < .001$), and significantly greater arousal in the Overload condition ($M = 3.08$, $SD = 1.163$) than the Boredom condition ($p < .001$). Please refer to Figure 4.

Discussion

In both EDA and SAM measures of arousal, the Boredom condition was reported to have generated the lowest levels of arousal. Both the Balance and the Overload condition featured significantly higher EDA response and self-reported arousal. This potentially points to a relationship between increased challenge and increased physiological and self-reported arousal.

This relationship is further suggested by the disparity between the EDA responses in the Balance condition and the Overload condition, in which Overload prompted a significantly higher EDA response than Balance. As such, physiological arousal is found to increase the more challenging the play experience.

These initial findings differ from those of Kneer et al., in which no effect on physiological arousal (measured by EDA) as a consequence of difficulty was found [6]. It is possible that the difficulty manipulations used in this study are incomparable to those used by Kneer et al. For example, the 'Overload' condition in this research was developed with the intention of overwhelming the player and rendering the task impossible; while Kneer et al. aimed for higher difficulty, it's possible it was not to this extreme. Additionally, play times in Kneer et al. were twenty minutes in length, whereas the play times used in this study were ten minutes; it's possible that participants in the study by Kneer et al. played long enough to habituate to the high difficulty.

An additional explanation may be found in research undertaken by Ravaja et al., in which increased EDA response was found when a player was killed or wounded, or killed or wounded an enemy [12]. As such, increased exposure to the death/wounding of both enemy opponents and the player-character in the Balance and especially the Overload condition may also be partially responsible for increased EDA response.

Interestingly, EDA results revealed arousal difference between Balance and Overload that the SAM did not. This may suggest that EDA is capable of finer granularity than the SAM measure. Alternatively, the decision to include the Overload condition last for all participants, and thereby avoid frustration contaminating data from conditions played afterward, opens the argument that the current results are partially influenced by EDA drift. Moving forward, we plan to collect a sample of data with fully counterbalanced conditions so as to allow us to identify the amount of influence physiological drift has and to control for it as needed. Finally, as EDA is a real-time measure whereas the SAM is taken post-play, it is possible that the participant's arousal had diminished since play to the point of no difference between post-play Balance and post-play Overload. As such, this would highlight the benefits of using a real-time measure for PX evaluation.

Conclusion and Future Work

This research represents initial findings from a larger program of research. A potential relationship between physiological arousal, challenge, and self-reported arousal has been explored. Ongoing research plans to further explore this relationship, with a larger sample size, as well as the relationship between additional physiological measurements and experiential constructs. Ongoing research aims to contribute to the formation of a unified approach to an understanding of psychophysiological evaluation of the PX.

Acknowledgements

We thank Valve Corporation for supporting this research by kindly providing game keys for participant compensation.

References

1. John L. Andreassi. 1995. *Human Behavior and Physiological Response* (3rd. ed.). L. Erlbraum Associates, Mahwah, NJ.
2. Regina Bernhaupt, Wijand Ijsselsteijn, Florian Mueller, Manfred Tscheligi, and Dennis Wixon. 2008. Evaluating User Experience in Games. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '08), 3905-3908. http://dl.acm.org/citation.cfm?doid=1358628.1358953
3. Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1: 49-59.
4. Johannes Keller, Herbert Bless, Frederik Blomann, and Dieter Kleinböh. 2011. Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology* 47, 4: 849-852.
5. J. Matias Kivikangas, Guillaume Chanel, Benjamin Cowley, Inger Ekman, Mikko Salminen, Simo Järvelä, and Niklas Ravaja. 2011. A review of the use of psychophysiological methods in game research. *Journal of Gaming & Virtual Worlds* 3, 3: 181-199.
6. Julia Kneer, Malte Elson, and Florian Knapp. 2016. Fight fire with rainbows: The effects of displayed violence, difficulty, and performance in digital games on affect, aggression, and physiological arousal. *Computers in Human Behavior* 54: 142-148.
7. Sohye Lim and Byron Reeves. 2009. Being in the Game: Effects of Avatar Choice and Point of View on Psychophysiological Responses During Play. *Media Psychology* 12, 4: 348-370.
8. Regan L. Mandryk, M. Stella Atkins, and Kori M. Inkpen. 2006. A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), 1027-1036. http://dl.acm.org/citation.cfm?doid=1124772.1124926
9. Lennart Nacke and Craig A. Lindley. 2008. Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (Future Play '08), 81-88. http://dl.acm.org/citation.cfm?doid=1496984.1496998
10. Lennart E. Nacke, Mark N. Grimshaw, and Craig A. Lindley. 2010. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers* 22, 5: 336-343.
11. Niklas Ravaja, Timo Saari, Mikko Salminen, Jari Laarni, and Kari Kallinen. 2006. Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation. *Media Psychology* 8, 4: 343-367.
12. Niklas Ravaja, Marko Turpeinen, Timo Saari, Sampsa Puttonen, and Liisa Keltikangas-Järvinen. 2008. The Psychophysiology of James Bond: Phasic Emotional Response to Violent Video Game Events. *Emotion* 8, 1: 114-120.
13. Carmen V. Russoniello, Kevin O'Brien, and Jennifer M. Parks. 2009. EEG, HRV and psychological correlates while playing Bejeweled II: A randomized controlled study. *Studies in Health Technology and Informatics* 144: 189-192.
14. Robert M. Stern, William J. Ray, and Karen S. Quigley. 2001. *Psychophysiological Recording* (2nd. ed.). Oxford University Press, New York, NY.