

Evaluation methods of exergames

Methods and instrument for assessing the quality of exergames

Seminar Paper by Dustin Grünwald

Reviewer: Dr.-Ing. (habil.) Stefan Göbel

Supervisor: Polona Caserman

Darmstadt



Electrical Engineering and
Information Technology
Department
Serious Games

Erklärung zur Abschlussarbeit gemäß § 22 Abs. 7 und § 23 Abs. 7 APB der TU Darmstadt

Hiermit versichere ich, Dustin Grünwald, die vorliegende Seminar Paper ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mir ist bekannt, dass im Fall eines Plagiats (§ 38 Abs. 2 APB) ein Täuschungsversuch vorliegt, der dazu führt, dass die Arbeit mit 5,0 bewertet und damit ein Prüfungsversuch verbraucht wird. Abschlussarbeiten dürfen nur einmal wiederholt werden.

Bei der abgegebenen Thesis stimmen die schriftliche und die zur Archivierung eingereichte elektronische Fassung gemäß § 23 Abs. 7 APB überein.

Bei einer Thesis des Fachbereichs Architektur entspricht die eingereichte elektronische Fassung dem vorgestellten Modell und den vorgelegten Plänen.

Darmstadt,

D. Grünwald



Contents

Abstract

Hint:

To be done

1 Introduction

1.1 Motivation

Fitness video games, now referred to as exergames, gained a lot of popularity in the last years, mostly due to the development of cheap and small motion sensors. This is not surprising as they enable to combine the healthiness of sports with the motivating aspects of video games. "There was moderate quality evidence that AVGs can result in benefits for self-esteem, increased energy expenditure, physical activity and reduced body mass index in children and adolescents who used AVGs in the home environment" [1] increased self-esteem in boys ($p < 0.05$; ES: 0.58; 95% CI: 0.08-1.07) and reduction of girls' mental confusion ($p < 0.05$; ES 0.58; 95% CI 0.15-1.06). **andrade2020effect** They can adress health problem of video gamers as they are at a higher risk of being overweight [2] or frequent online game players who report exercising at a lower frequency [3]. Also elder people with less connections to video games can train their balance using exergames [4]. Playing Pokémon Go was shown to be associated with statistically significant number of daily steps [5]. This becomes more impressive considering that it is the third most popular smartphone game of all time [6]. However, there are also studies coming to the conclusion that offering exergames, instead of regular video games, to children does not increase their physical activity under naturalistic circumstances [7]. To further improve exergames, reliable ways to evaluate them are advantageous. The available research mostly focused on pure videos games and serious games. Serious games are all video games which pursue another "characterization goal" aside from entertainment, e.g. fighting cancer cells in Re-Mission [8] was shown to help children with cancer to develop a more positive attitude towards their illness with positive health aspects, as shown in randomized controlled trials [9]. Exergames represent a subgenre of serious games with a focus on excersing i.e. body movements. As serious games benefit, due to their double mission, from specialiced evaluation methods [10] in comparison to general video games, exergames again can benefit from evaluation methods that consider the involved physical activity. This paper will give insights to the challenges and possibilities.

1.2 Approach

Three studies were proposed as interesting by the supervisor Polona Caserman who is an researcher in the field of serious games They acted as a good starting point for snowball sampling. Other studies were often found using ChatGPT 3.5 with queries in the format of "list some studies about ..." which usually gave a good overview. Although the name of the listed studies were wrong many times, the author name and the release year could be used to find the related paper in Google or Google Scholar. E.g. the query "list some studies about to evaluation of exergames" led to these studies [11] [4]. To find evidence for several trains of thought or to answer upcoming questions, key words where put into Google Scholar, e.g. "Heart rate variability emotion" to find more about the connection of heart rate variability and emotions.

2 Measuring outcomes

"The [metastudy] results show that playing AVGs significantly increased heart rate, VO2 (oxygen consumption), and EE (energy expenditure) from resting" [12]

2.1 Trainingsdauer

Messungen: "6-minute walk test (6MWT), and secondary outcomes were body mass, self-reported physical activity, sleep quality, Berg Balance Scale, Short Physical Performance Battery, fast gait speed, dynamic balance, heart rate recovery after step test, and 6 cognitive tests." Anzahl Trainings pro Woche: "no (1×), most (2×), and medium (3×/week) effects of exergaming volume in most motor and selected cognitive outcomes." [13]

3 Impact of sport on physiological measures

"HRV measurements change with exercise intensity. For example, Hottenrott et al. (21) demonstrated that increased exercise intensity was associated with an increased LF/HF ratio, reflecting increased sympathetic tone in recreational cyclists. Elite triathletes challenged with moderate-intensity and high-intensity exercise tests showed increased sympathetic and parasympathetic tone, respectively (14). Similarly, during exhaustive exercise, Pichon et al. (34) reported increases in parasympathetic input. These studies indicate that, depending on the length and intensity of exercise, the autonomic input may be sympathetic up to a point with parasympathetic predominance at exhaustion." SUMMARY: ShortTerm exercise => increased LF/HF ratio => Sympathetic, LongTerm bei exhaustion => Parasympathetic; "Section Summary: HRV Changes are Detectable in Overtraining and the Over-Reaching (PreOverTrained) State. Multiple studies illustrate that HRV variables change in the context of overtraining (summarized in Table 5). Several studies reported changes in HRV variables that reflect enhanced sympathetic activity at rest in the overtrained state while others identified a predominance of parasympathetic activity at rest in the overtrained states. The different forms of overtraining that have been described may explain the changes. Specifically, overtraining associated with increased exercise volume (sets x repetitions) and intensity (percent repetition maximum) has been associated with a sympathetic or parasympathetic predominance, respectively (21). Recent studies also provide evidence that in the early stages of overtraining, designated as overreaching phases, HRV may be characterized by a sympathetic dominance. When this over-reaching phase occurs over a longer period of time, then, the development to that of the overtrained state, as characterized by parasympathetic predominance, may ensue" SUMMARY: Overtraining => Sympathetic, then when exhausted Parasympathetic dominance; As summarized by Makivic, during exercise, HRV typically shifts toward a higher LF/HF ratio, reflecting increased sympathetic activation and after exhausting exercise, towards a more parasympathetic activation. [14]

HRV may become less informative during exercise or periods of increased heart rate as root mean square of successive differences and approximate entropy of R-R intervals (ApEnRR), did not significantly differ during exercise[15]

Physical activity leads to increased heart rate **hammond 1985normal**, [15], [16], vigorous activity leads to rapid breathing and substantial increase in heart rate. [16]

3.1 Smartwatches

Energy expenditure: Fitbit Surge HR and TomTom watch($r = 0.62-0.69$) average HR: Apple Watch, Fitbit Surge HR, and TomTom watch($r = 0.47-0.74$) Peak HR: all smartwatches ($r = 0.59-0.65$). [58 Pope 2019 Validation of four smartwatches in EE and HR]

4 Existing evaluation methods

There are multiple methods available to evaluate games with different characteristics. The methods can be applied during or after the gaming sessions. They can be subjectively, measured by a human and often the players themselves, or objectively using a device. These distinctions create the following table.

4.1 Questionnaires

One important part to understand player's impressions towards video games are questionnaires as they can gather large amounts of data in a standardized way. One major drawback is that they need some time to complete and thus, are often conducted some time after the game was played. The Game Experience Questionnaire (GEQ) tries to solve this problem using a more concise In-game questionnaire [17].

Another validation study use exploratory and confirmatory factor analyses to find partial support for the GEQ and large support for the PENS [18], also proposing some revisions to the factor structures GEQ: negative affect, tension/annoyance and challenge => negativity; intuitive controls, competence => competence PENS: 3 categories presence, autonomy and relatedness supported

4.1.1 The Player Experience of Need Satisfaction (PENS) Model

The creators of this questionnaire claim that fun and satisfaction are outcomes of psychological processes and not the processes themselves [19]. So in order to create an entertaining game for the player, you need to understand the underlying processes and be able to describe the "underlying energy that fuels actions" [19]. Thereby, PENS is grounded on the well-established self-determination theory from the 1980s. It elicited validated measures in many fields [20]–[22]. This theory suggests that there are three basic psychological needs [23]:

- Autonomy: The need to feel in control of one's actions and have the freedom to make choices.
- Competence: The need to feel capable, effective, and skilled in one's pursuits or endeavors.

	Live	Afterwards
Subjective		
- Self-reported	Live questionnaires, think-aloud	Questionnaires, interviews, focus groups, heuristic evaluation
- Observed	Behavioral observation	
Objective	Physiological measures, eye tracker, game metrics	

- Relatedness: The need to feel connected, supported, and engaged in meaningful relationships with others.
- -
- Presence/Immersion: emotional engagement in the game
- Intuitive Controls

4.1.2 Game Experience Questionnaire (GEQ)

The Game Experience Questionnaire was developed 2007 by a European research project named FUGA [24]. It is widely used in at least 515 papers [24]. Due to its importance, its story shall be presented. To create the questionnaire, FUGA first formed focus groups of both frequent and infrequent gamers with three questions to come up with reasons for, feelings when and feelings after playing video games. Then, five experts including psychologists and experts on measurement development met. They used theoretical considerations as well as the focus group's results to create a model to understand gaming experience. They came up with ten factors that were used to create a questionnaire with 92 items and a 5-point Likert-scale from "not at all" (0) to "extremely" (4). Then, 380 participants were recruited online to test the questionnaire with a game of their choice. Then a exploratory factor analysis was run on the results. They came up with a 7-factor solution with 82 items. It explained 52 % of variance in all items. Most questions have more than 30 % correlation to only one factor. The seven factors are:

1. Immersion (previously two factors Sensory and Imaginary Immersion)
2. Tension (said to unforeseenly emerged from Negative Affect although the factor Suspense already existed)
3. Competence (includes previous factor Experienced Control)
4. Flow
5. Negative Affect
6. Positive affect
7. Challenge

The factors "Connectedness" and "Negative affective experiences related to playing with others" were shifted to an own questionnaire, the "social presence module" including another new category "behavioural involvement"

Systematic Review and validation of the GEQ

A paper from 2018 did a systematic review using 73 studies that applied the GEQ [24]. Many critic points were found for both a GEQ itself but also for the studies that used it

- Only 31 papers of 73 offered an explanation why the GEQ was used (31 because it is validated, 10 due to popularity, 8 due to multidimensional structure and 6 for being theoretically and empirically founded)
- 47 Papers did not state the number of questions used whiel the GEQ offers a 33- and a 42-items version
- Often times, papers used a selection of factors or items or adapted the items with mostly missing reasoning, partially explained by overlapping after using factor analysis

- The used scale was not reported by 40 papers, 27 did not use the origin scale and 3 of them even used a mix of 5 and 6 answer options instead of 5
- One main critic point was that there are two versions of the GEQ, from 2007 and 2013 with none of them formally published. The 73 analyzed papers referenced the 2007 version only once and the 2013 version not a single time. Often times, the GEQ was cited as "Manuscript in preparation"
- 17 papers stated Cronbach's α for internal consistency of the factors with following values:
 - Flow and Competence: 0.7 to 0.94
 - Positive Affect and Immersion: 0.49 to 0.85
 - Negative Affect, Challenge and Tension: 0.3 to 0.74

12 of these 17 papers stated low internal consistency, especially for challenge and negative affect. No paper could reproduce the 7-factor structure but six proposed using six factors with Negative and Positive Affect being summarized as "joy".

Finally, a validation study was conducted with 633 participants in an online survey (age = 33.47 ± 10.57 , game experience = 19.5 ± 8.9 years) using the 33-items GEQ. A confirmatory factor analysis resulted in the factor Negative Affect not being satisfactory due to a low reliability coefficient and the factor Challenge just barely. Also, an exploratory factor analysis was conducted. It came to the conclusion to remove some items that did not load well on the given or any other factor. Flow may be renamed to "loss of time". The Challenge factor only has 3 remaining items with need for filling up. Also Tension and Negative Emotion should be combined to one factor which results in these factors:

1. Immersion
2. Competence
3. Flow
4. Negative Affect
5. Positive affect
6. Challenge

Unfortunately, the reviewing paper does not address multiple issues. What are possible reasons why the results of the conducted study differ from the original study? Are there maybe missing questions that could lead to a new factor that the GEQ is missing? Can the items of Negative and Positive Affect be summarized like in six reviewed papers? (with inverting the questions for Negative Affect) Can the factor Challenge be added to the factor Competence with inverted questions?

4.1.3 Flow State Scale

Hint:

More sources to be included

Flow describes a positively experienced state, where a person is totally in a task with challenge and skill level holding a perfect balance [25]. Further research has shown that, when experiencing flow, the perceived challenge level is even slightly higher than the skill level (8.3 vs 7.3) [26]. Getting in flow state would be pleasant for exergames too as people seem to show peak performance while in flow state [26]. It also plays a role in motivation and enjoyment [26]. To evaluate flow, there is a questionnaire called the "Flow State Scale". It originally consisted of 54 items with 9 factors but could be reduced to 36 items while keeping a high reliability with coefficient alphas rated above 0.80 [26]. The 9 original factors are [26]:

1. Sense of Control
2. Challenge-Skill-Balance
3. Clear Goals
4. Merging with task
5. Unambiguous feedback
6. Concentration
7. Loss of Self-Consciousness
8. Transformation of Time (potential for removal as weakest factor)
9. Intrinsic motivation (potential for removal as it may be more of a higher-order factor)

4.1.4 Player Experience of Need Satisfaction (PENS)

[Hint:](#)

[To be done](#)

4.2 Psychophysiological measures

4.2.1 applicability in Videos games

"Some of the earliest are from Mandryk and Inkpen (2004) and Hazlett (2006), who have presented studies (albeit with small sample sizes) supporting the use of psychophysiological measures in game research. More recently, Nacke (2009) and others have published studies as an attempt at a common methodology for a design-oriented approach. Their papers evaluate EEG (Nacke et al. 2010b), EDA, HR (Nacke 2009; Drachen et al. 2009) and facial EMG (Nacke 2011). **kivikangas2011review** "clearly showed the necessity of proper experimental design and that care must be taken in interpreting the signals: otherwise, for instance, self-reported and physiologically indexed emotions may turn out to assess significantly different things" **kivikangas2011review** "All these studies demonstrate that physiological signals are closely related to players' self-reported emotional states and behaviour, while Chaneel and others (2011) found support that the fusion of several physiological

modalities increases the recognition accuracy. This shows that multiple measurements are still needed for a reliable interpretation of the player experience" **kivikangas2011review**

ToDo: Paper 54 lesen

Psychophysiological measures can be a helpful way in evaluating exergames because they " provide an objective, continuous, real-time, noninvasive, precise and sensitive way to assess the game experience" **kivikangas2011review**. "A large number of studies have shown that psychophysiological measures can be used to index emotional, motivational and cognitive responses to media messages" **kivikangas2011review** "physiological signals could be used, for example, at the player test phase to identify strong emotional episodes and compare them to expectations, or to control the successful emotional elicitation of game event designed to be emotionally arousing" **kivikangas2011review** "Psychophysiological research is defined as using physiological signals to study psychological phenomena (Cacioppo et al. 2007: 5)." **kivikangas2011review** The human body offers many ways to gain insights into someone's emotions state, e.g. heart rate and skin conductance together can be used to effectively recognize some basic emotions [27]. Also, these measurements can be taken continuously during the gameplay and can be mapped to a recorded gameplay session afterwards [28] which allows great post-analysis. Nonetheless, they should not be relied on alone. The researcher should always ensure the correct interpretation of the signals [28], e.g. by validating them using other measurement methods like questionnaires One major drawback of these measures is that they are not only be affected by the game experience itself. Any movement with the body affects physiological measures, e.g. physical exercise results in heart rate increase [29]. This can be obviously a problem in evaluating exergames where the body response usually consists of a psychological and a physical reaction. Aside from moving, also any stimulus in the environment that is not gameplay-related can be disturbing and affect the measurement [28]. Also the demographic background of any person and the baseline of the measured variable in idle state should be taken into consideration [28] Although it is often hard to directly map physiological measurements to basic emotions, Russell's circumplex model can be applied to simplify things. It uses just two dimensions valence ("pleasant/unpleasant emotion") and arousal ("low/high stimulation"). Positive and negative emotional valence can be assessed by using facial EMG **kivikangas2011review**. They can be more easily measured and concrete emotions can be derived from them [30] This is why many measuring methods are often not applicable for exergames which will be handled in the next chapter Another important point for exergames is portability and applicability during exercise. There may be games with heavy movement where a small or wireless measurement device may be advantageous Now comes an overview of measure possibilities. For more details, the just mentioned paper by Nacke 2014 is recommended

4.2.2 Facial Electromyography (EMG)

Sensors will be attached to the skin and measure the electric and such physical activity of muscles. They can be used to detect facial expressions. It was shown that eye [31] and eyebrow muscles are a good indicator for positive and cheek muscles a good indicator for negative emotions [28], [32]. Disadvantages are discomfort wearing face sensors, the inability to talk during gameplay and unnatural facial expressions due to feeling the sensor [28] Also computer vision techniques can be used to obtain facial expressions but when discovering smiling, "The results showed that EMG has the advantage of being able to identify covert behavior not available through vision. Moreover, CV appears to be able to identify visible dynamic features that human judges cannot account for" [33]

4.2.3 Electrodermal Activity (EDA) / Skin conductance

Two electrodes measure the conductivity of the skin. This is correlated to the sweat gland activity and that again is linked to emotional arousal [28], [34]. The measuring device can function wireless and be made out of soft materials [35] do not be disturbing during exercise The main disadvantage is some seconds of latency which makes it more difficult to link a signal to a cause [28]

"Significant covariation was obtained between (a) facial expression and affective valence judgments and (b) skin conductance magnitude and arousal ratings" [36]

4.2.4 Cardiac activity / Heart Rate

Heart rate was shown to effectively recognise all four basic negative emotions anger, fear, disgust and sadness [37] and is tied to emotional arousal [28]. There are many possibilities to measure heart rates. Chest strap monitors are very precise and are affordable (75 €) [38] and also current smartwatches can deliver good results **gilinov2017variable**.

4.2.5 Heart Rate Variability

The variability of the heart rate can be also obtained from measuring the heart rate but needs further analysis [28]. There is emerging analysis for its role in regulated emotional responding [39] with higher variability typically correlating with better emotional regulation [source?]. However, it was shown that heart rate variability alone is not that effective in recognizing emotions with a baseline below 50 % [40].

4.2.6 Electroencephalography (EEG)

Electroencephalography is used to measure brain waves with a high temporal accuracy [28]. The activity of different brain regions can be visualized but there is a lot of interpretation possibilities, especially since the origin of the brainwaves is not apparent [28]. At least it can be used to measure a player's engagement in a video game, e.g. there was increased activity when doing in. Super Meat Boy [41]. With the Emotive EEG headset [42], there is also portable technology available. "Relatively greater left frontal activity indicates a propensity to approach a stimulus, whereas relatively greater right frontal activity indicates a propensity to withdraw from a stimulus (Davidson 1998). It must be emphasized that frontal asymmetry is not a measure of positive or negative affect per se, but it taps a broader motivational tendency towards approach-related or withdrawal-related behaviours and emotions (Allen et al. 2001; Davidson 1998)" **kivikangas2011review**

4.2.7 Respiration rate

Studies reported higher respiration rate during arousal [43], increases in inspiratory time when laughing and decreases during disgust **boiten1998the**. Although this measure does not give as much insights as others, it can be well estimated from using heart rate sensors, so no additional hardware is necessary when heart rate is already measured. The root mean square error then is 0.648 per minute [44].

4.2.8 Eyes

"The best accuracy achieved by fuzzy integral fusion strategy is 87.59%, whereas the accuracies of solely using eye movements and EEG data are 77.80% and 78.51%, respectively." to detect "positive, negative and neutral" emotions. Uses 33 different features just from the eyes (pupil diameter, dispersion, fixation, blinking, saccade (ruckartiges Verhalten)) [45]

4.2.9 Body

"To extend from psychophysiological measurements, there is some evidence that body movement and position (measured by acceleration sensors or position cameras) might be associated with attention, interest and emotions (FUGA Contributors 2009; Kivikangas and Ravaja in preparation)." **kivikangas2011review**]

4.3 Comparing exergames with conventional interventions


The effectiveness of exergames can be evaluated by comparing their outcomes with conventional interventions. Many evaluations methods and dependant variables are conceivable

Physical activity in inactive children

A study from 2010 has shown the big potential of exergames for a physical education classroom [11]. By the teacher's observation, four 5th grade class children did not move as much as other children. These "inactive" labeled children should be motivated to move more and participated in the study. Contrary to the default lesson where children played e.g. ball games, the exergame lesson consisted of 11 exergame stations. There, the children rotated every 10 minutes and played games like "Fit Interactive 3 Kick", a martial art simulator that uses three foam pads to be hit. The measured variable was "physical activity" which was defined as moving a large muscle group. It was shown that the inactive children increased their physical activity from 16 min to 92 min during a lesson. This massive increase was only made possible by enhancing the time of opportunities for participating in exercises from 38 to 116 min. This was mainly reasoned by the fact that the teacher spent way less time providing instructions. It is unclear if the effects would persist longer than the given six weeks. Additionally, the study relies on one single physiological measure as it spared measuring the heart rate in favor of saving time during the lessons. Still, the pure method of comparing exergames with a default setting enabled interesting insights and should be further examined.

Balance training for elder people

A study from 2012 evaluated the effectiveness of exergame exercises on balance improvement in older adults [4]. After six weeks, significant improvements were observed in typical balance tests such as the Berg Balance Scale. The positive effects partially persisted even after 6 weeks without further exercises. The exergame approach showed advantages in terms of feasibility and attractiveness compared to conventional exercises, due to instant and positive feedback and the more challenging nature. The study suggests the need for further research to examine the long-term effects and compare the exergame approach with traditional physical therapy.



Riding a bike while watching a bike tour video that gets played faster the faster you cycle was compared with riding a bike to play a video game with way more motivation for the video game. [46]

4.4 Evaluation of excersises



5 Evaluation methods for exergames



List of Acronyms

Bibliography

- [1] I. K. dos Santos, R. C. da Silva Cunha de Medeiros, J. A. de Medeiros, P. F. de Almeida-Neto, D. C. S. de Sena, R. N. Cobucci, R. S. Oliveira, B. G. de Araújo Tinoco Cabral, and P. M. S. Dantas, “Active video games for improving mental health and physical fitness - an alternative for children and adolescents during social isolation; an overview”, 2020.
- [2] Melchior, “Internet and video game use in relation to overweight in young adults”, 2014.
- [3] Kowert, “Unpopular, overweight, and socially inept; reconsidering the stereotype of online gamers”, 2014.
- [4] Lai, “Effects of interactive video-game based system exercise on the balance of the”, 2013.
- [5] Khamzina, “Impact of pokemon go on physical activity”, 2019.
- [6] https://en.wikipedia.org/wiki/List_of_most-played_mobile_games_by_player_count, [Online; accessed 31-May-2023].
- [7] Baranowski, “Impact of an active video game on healthy children’s physical activity”, 2012.
- [8] <https://www.re-mission2.org/>, "[Online; accessed 31-May-2023]".
- [9] Kato, “A video game improves behavioral outcomes in adolescents and young adults with cancer a randomized trial”, 2008.
- [10] P. Caserman, K. Hoffmann, P. Müller, M. Schaub, K. Straßburg, J. Wiemeyer, R. Bruder, and S. Göbel, “Quality Criteria² for Serious Games: Serious Part, Game Part, and Balance”, *JMIR Serious Games*, vol. 8, no. 3, Apr. 2020. DOI: 10.2196/19037.
- [11] Fogel, “Effects of exergaming on physical activity among inactive children”, 2010.
- [12] Peng, “Is playing exergames really exercising a meta-analysis of energy expenditure in active video games”, 2011.
- [13] Hortobagyi, “Effects of exercise dose and detraining duration of mobility at late midlife; a randomized clinical trial”, 2021.
- [14] Makivic, “Heart rate variability (hrv) as a tool for diagnostic and monitoring performance in sport and physical activities”, 2013.
- [15] S. BOETTGER, C. PUTA, V. K. YERAGANI, L. DONATH, H.-J. MÜLLER, H. H. W. GABRIEL, and K.-J. BÄR, “Heart rate variability, qt variability, and electrodermal activity during exercise”, 2010.
- [16] Haskell, “Updated recommendation for adults from the american college of sport midicine and the american heart association”, 2007.
- [17] Poels, “Game experience questionnaire”, 2007.
- [18] D. Johnson, M. J. Gardner, and R. Perry, “Validation of two game experience scales: The player experience of need satisfaction (pens) and game experience questionnaire (geq)”, 2018.
- [19] Rigby, “The player experience of need satisfaction (pens)”, 2007.

-
- [20] G. Pietrabissa, A. Rossi, M. Borrello, G. M. Manzoni, S. Mannarini, G. Castelnuovo, and E. Molinari, "Development and validation of a self-determination theory-based measure of motivation to exercise and diet in children", 2020.
- [21] J. Lohmann, A. Souares, J. Tiendrebéogo, N. Houlfort, P. J. Robyn, S. M. A. Somda, and M. D. Allegri, "Measuring health workers' motivation composition validation of a scale based on self-determination theory in burkina faso", 2017.
- [22] D. K. Richards, M. R. Pearson, and C. A. Field, "Further validation of the treatment self-regulation questionnaire for assessing motivations for responsible drinking; a test of self-determination theory", 2021.
- [23] E. L. Deci and R. M. Ryan, "Intrinsic motivation and self-determination in human behavior", 1985.
- [24] Law, "Systematic Review and Validation of the Game Experience Questionnaire (GEQ) - Implications for Citation and Reporting Practice", *JMIR Serious Games*, vol. 8, no. 3, Apr. 2018. doi: 10.1234.
- [25] Jackson, "Psychological correlates of flow in sport", 1998.
- [26] Jackson, "Development and validation of a scale to measure optimal experience the flow state scale", 1996.
- [27] Hamdi, "Emotion recognition based on heart rate and skin conductance v6", 2015.
- [28] Nacke, "Physiological game evaluation", 2015.
- [29] Javorka, "Heart rate recovery after exercise", 2002.
- [30] Seo, "Automatic emotion-based music classification for supporting intelligent iot applications", 2019.
- [31] Ravaja, "Phasic emotional reactions to video game events", 2018.
- [32] Mandryk, "Using psychophysiological techniques to measure entertainment", 2006.
- [33] M. Perusquía-Hernández, S. Ayabe-Kanamura, K. Suzuki, and S. Kumano, "The invisible potential of facial electromyography", 2019.
- [34] M. E. Dawson, A. M. Schell, and D. L. Fillion, "The electrodermal system; in: J. t. cacioppo, l. g. tassinari and g. g. berntson (eds), handbook of psychophysiology, new york, ny: Cambridge university press", 2000.
- [35] Kim, "Soft wireless bioelectronics and differential electrodermal activity for home sleep monitoring", 2021.
- [36] P. J. LANG, M. K. CREENWALD, M. M. BRADLEY, and A. O. HAMM, "Looking at pictures: Affective facial visceral and behavioral reactions", 1993.
- [37] Levenson, "Blood sweat and fears - the automatic architecture of emotion", 2003.
- [38] <https://www.amazon.de/Polar-Herzfrequenz-Sensor-Bluetooth-Wasserdichter-Brustgurt/dp/B07PM54P4N>, "[Online; accessed 31-May-2023]".
- [39] Appelhans, "Heart rate variability as index of regulated emotional responding", 2006.
- [40] Ferdinando, "Emotion recognition by heart rate variability", 2014.
- [41] McMahan, "Evaluating electroencephalography engagement indices during video game play", 2018.
- [42] <https://www.emotiv.com/epoc-flex/>, [Online; accessed 31-May-2023].
- [43] Homma, "Breathing rhythms and emotions", 2008.
- [44] Natarajan, "Measurement of respiratory rate using wearable devices and applications to covid-19 detection", 2021.

-
- [45] Y. Lu, W.-L. Zheng, B. Li, and B.-L. Lu, “Combining eye movements and eeg to enhance emotion recognition”, 2015.
 - [46] Hardy, “Adaption model for indoor exergames”, 2011.



A Appendix

The appendix goes here... english