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] There are hints on increased self-esteem in boys and reduction of girls' mental confusion [2]. They can adress health problem of video gamers as they are at a higher risk of being overweight [3] or frequent online game players who report exercising at a lower frequency [4]. Also elder people with less connections to video games can train their balance using exergames [5]. Playing Pokémon Go was shown to be associated with statistically significant number of daily steps [6]. This becomes more impressive considering that it is the third most popular smartphone game of all time [7]. 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 [8]. 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 [9] 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 [10]. Exergames represent a subgenre of serious games with a focus on excersing ie body movements. In comparison to general video games, serious games benefit, due to their double mission, from evaluation methods that consider both serious part and game part plus their balance [11]. Exergames again can benefit from evaluation methods that consider the involved physical activity. This paper will give insights to the possibilities and challenges. "moderate-intensity physical exercise is good at improving emotional response to negative stimulations" [67] Lack of physical activity is a significant contributing factor to preventable deaths in the American population [12].

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

[13] [5]. 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.

"heart rate (exergame OR exercise video game OR active video game)" "Games User Research" "video game shooter exertion motion controlled" Furter more, this paper serves as an overview. To keep concise, when summarizing other papers, most often only key information are included. Terms like "may" and "possible" will hint on possible limitations of the studies referred to. The reader is encouraged to dive deeper into topics of their interest.

1.3. Target

The target of this paper is to be able to examine how well a exergame works. If such a game works, can be supported by scientific studies or by winning game wards [11] but also by playtime. In exergames, you want the player a) to feel pleasure in the moment and b) reach the desired training effects. Pleasure in exergames can be measured in many ways including player assessments, psychophysiological measures or general game metrics like play duration and frequence. Training effects can be measured best based on outcomes including anthropometric (weight, height) and physiological measures, fitness tests, daily metabolic rates and much more. In order to maximize training effects, regular exercise is advantageous, e.g. as shown by Heaselgrave for muscle growth [14]. Also appropriate session durations are needed. This paper will give an overview how pleasure and training effects were measured in the scientific research and what other opportunities are there. The research specifically on the evaluation of exergames will be included as far as it is available. Otherwise, the aspects of video games and physical excersises are examined mostly individually and consolidated later.

1.4. Terminology of exergame

There a multiple variations to the term exergame including As occuring in the title, this papers will use the term exergame as its widely used in literature, short and catchy [15]. Exergame is a blend word compositing of exercise and (video) game. In the field of scientific research, there is common ground that exergames are video games. Another aspect is that the games need to involve some sort of exercise or physical acitivity. The difference is made clear by Caspersen who defined [16] physical activity in 1985 as "any bodily movement produced by skeletal muscles that results in energy expenditure" and exercise as "subset of physical activity that is planned, structured, and repetitive and has as a final or an intermediate objective the improvement or maintenance of physical fitness". So, exercise is the more narrow definition that additionally requires some plan, structure and repitition. Some authors prefer this narrow definition for exergames while most researchers use the broader definition including any physical activity [15]. This paper will affiliate and also use the broader definition. That means, games like Pokemon Go will be count as exergame since you engage a lot in moving although you don't do it in a planned way.

2. Measuring outcomes

"The [metastudy] results show that playing AVGs significantly increased heart rate, VO2 (oxygen consumption), and EE (energy expenditure) from resting" [17]

3. Motion detection

Full-body motion need to be recognized correctly in order to ensure the desired execution of physical activities [18]. When using head mounted displays, a low latency is required to enhance the visual feedback and reduce cybersickness [18].

4. Sonstiges

4.1. Avatars

When players of Nintendo Wii Fit used their ideal self as avatars instead of their current selfs, Jin [19] showed that they experienced greater perceived interactivity.

4.2. Impact of training frequency of exergames

One exercise per week led to no effects and three per week to medium effects while 2 exercises per week resulted in most effects in most motor and selected cognitive outcomes. These outcomes wer measured by: 6-minute walk test (6MWT), 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. [20]

4.3. Gameplay metrics

Gameplay metrics analysis in behavior analysis is a valuable tool for tracking and analyzing user behavior in complex computer games [21]. It complements existing methods, provides quantitative data, and enhances understanding of the play experience. Despite challenges, it is widely recognized in the game industry and academia as a valuable approach for evaluating and testing games. Drachen identified three categories: Generic gameplay metrics (eg durations, sessions, location, played part of the game, progress), Genre specific gameplay metrics (eg character movement, interactions, interface usage) and Game specific gameplay metrics. Marston et al [22] tried integrating the genre exergames in a video game genre map developed by Fencott in 2012 [23], with connections to Sports, (People) Simulation, Puzzle and Turn Based genres. Still, any genre for exergames can be imaginable,

Wols et al [24] showed 2018 that in-game play behaviour in MindLight, a game teaching relaxation methods, can predict improvements in anxiety symptoms 3 month later.

In 2011, Kivikangas et al [25] presented a tool that could record and map ingame events, ingame recordings, psychophysiological measures and live player feedback, which was also enabled by the tool. It can only be used in the Source engine by Valve Corporation.

4.4. Impact of sport on psychology

Anxiety reduction was shown after 20-minute treadmill run for experienced but not for inexperienced runners. [26], [27] Exercise can lead to increased beta-endorphin levels causing an analgesic and "feel-better" effect [26]. Also a muscle relaxation response could be found [26], [28]. Physiological arousal of people with higher aerobic power (the ability of the muscles to produce energy using oxygen) is lower and recovers faster to base level which also effects emotions: People that were made angry in experiments, excercised and then sat for some time, showed more intensive emotions, ie aggresive behaviour, at lower fitness levels [29]. Contrary to meditation and music appreciation programs, at long term, exercise led to a faster recovery the electrodermal response after applying different kinds of psychosocial stress [30] while additionally increasing physical fitness. Still, there are also many studies that did not come to the same conclusion which may also arise from habituation due to using the same social stressors [26]. A systematic review concluded that extrinsic motivation may be more influential in motivating individuals to start exercise in the short term, while intrinsic motivation may be more predictive in long-term.[31] "Regular exercise is associated with emotional resilience to acute stress in healthy adults" [32].

When people feel motivated by their own choices and have their basic psychological needs met, they are more likely to have positive emotional experiences during exercise. However, for individuals who are externally regulated, perceiving the intensity of the exercise as high can help sustain a better emotional response. [33] One study proposed that changes in emotions, stress, and effort would be associated with the intensity of running, with high intensity conditions causing more change and being perceived as more stressful and effortful. This hypothesis was supported by the results from the 1.7 km experiment but not the 5.0 km experiment [34]. As exergames rather use low intensities, prolonged sessions could be advantageous.

One study compared the effects of Wii exergames and traditional exercise on subthreshold depression in older adults, finding that exergames had a direct negative effect on subthreshold depression, with positive emotions mediating this effect. Self-efficacy was not found to be a significant mediator. These findings suggest that exergames may be a valuable tool in reducing subthreshold depression among older adults and could be considered in its treatment. [35]

4.5. Impact of sport on psychophysiological measures

4.5.1. HRV

As summarized by Makivic, during exercise, HRV typically shifts toward a higher LF/HF ratio, reflecting increased sympathetic activation and after exhausting excerise, towards a more parasympathetic activation. [36]

Sympathethic activity was shown to increase Electrodermal Activity but not the RMSSD parameter of HRV.[37]

4.5.2. EEG

Brain waves measured by EEG showed less latency after exercise-induced changes. [38] Also alpha brain waves showed 30-40 min increased magnitude after exhaustive tradmil walking which could be interpretes as an increased lack of attentiveness to the enivironmental surround [26].

Physical activity leads to increased heart rate [37], [39], [40], vigorous activity leads to rapid breathing and substantial increase in heart rate [40]. This could make the emotion-caused fluctuations harder to detect although research on this area is rare.

4.6. Smartwatches

Energy expenditure: Fitbit Surge HR and TomTom watch(r = 0.62-0.69) average HR: Apple Watch, Fitbit Surge HR, and TomTom watchr (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]

5. 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 distictions create the following table.

5.1. Think-aloud protocol

More frequent players talked much more about their insights and game strategies [41]. Tan et al examined the combination of think-aloud with physiological data. Many interesting physiological responses were not reflected by the think-aloud method and sometimes participants reported interesting experiences that could not me seen in the measures eg confusion.

5.2. Heuristics

Heuristics can be used as a help for developers during the early development [42]. Based on the Heuristic Evaluation for Playability (HEP) that shows some correlation to user experience [43] using 43 items, the Principles of Game Playability (PLAY) was developed. It has yo/no- items in 3 categories Game Play, Emotional Immersion and Usability & Game Mechanics. Strååt et al [44] applied the Herzberg's Two-Factor Theory to label items as motivational or hygienic ie if they are of a intrinsic or extrinsic nature or if they are creating positive feelings or disabling negative feelings [45]. If their impact is neutral to positive or neutral to negative. A group of 23 participants was used for sorting the items. Heuristics about usability were mostly hygiene factors eg "Players feel in control" and motivator were topics like storyline and immersion eg "The game is balanced with multiple ways to win". All in all, this two-factor-theory is a good reminder to think of certain aspects that effect game quality are more prone to have a positive and a negative impact.

5.3. 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 [46].

5.3.1. 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 [47]. So in order to create an entertaining game for the player, you need to unterstand the underlying processes und be able to describe the "underlying energy that fuels actions" [47]. Thereby, PENS is grounded on the well-established self-determination theory from the 1980s. Ut elicited validated measures in many fields [48]–[50]. This theory suggests that there are three basic psychological needs [51]:

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

Interestlingly, two of the three main factors, competence and relatedness, seem to be associated with positive emotions in long-term sport [52] (like novelty seeking and physical exertion).

A study from 2018 found large support for the PENS [53] and its categories presence, autonomy and relatedness. It proposed conclusion of Intuitive Controls into the Competence factor, especially if the players are more experienced with the game.

"the literature provides good evidence for the value of SDT in understanding exercise behavior". " intrinsic motives (e.g., challenge, affiliation, enjoyment) were positively associated with exercise behavior". "one study/sample performing correlational analysis to explore the links between health motives and exercise". "As expected from theory, controlled motives (social recognition, appearance/weight) did not predict, or negatively predicted, exercise participation". [31]. The disadvantage of high organizational effort that often goes with exercises [54] are not such a big deal in exergames.

5.3.2. Cronbach's alpha)

Internal consistency among one factor, e.g. 3 items (average intercorrelation among the three items) -1 to 1

5.3.3. composite reliability (CR))

Internal consistency among all factors, e.g. 10 factors and 30 items. Expected to be higher than Cronbach's alpha -1 to 1

5.3.4. Pearson correlation coefficient

linear correlation between two sets of data / normalized measurement of the covariance. -1 to 1

5.3.5. Discriminant validity

shows you that two tests that are not supposed to be related are, in fact, unrelated

5.3.6. The Player Experience Inventory (PXI)

Abeele et al created the Player Experience Inventory (PXI) to measure, analyse and understan a player's experiences [55]. It is split in immediate experiences based on game design choices and emotional experiences like immersion and mastery. Items were examined with the use of 64 Games User Research (GUR) experts, validated and evaluated with studies on totally 529 participants. They came up with 10 factors with 3 items each: Meaning, mastery, immersion, autonomy, curiosity, ease of control, challenge, progress feedback, audiovisual appeal and goals & rules. The inventory was set in relation with the PENS and the AttrakDiff, a questionnaire to measure hedonistic and pragmatic quality in a broder area and not just for games [56]. Here are the Pearson correlation coefficients:

Construct PXI	Construct PENS or AD	Pearson's R	Pearson'S R
Meaning	AD Attractiveness	0.638**	0.436**
Mastery	PENS Competence	0.884**	0.785**
Immersion	PENS Presence	0.596**	0.543**
Autonomy	PENS Autonomy	0.720**	0.603**
Curiosity	PENS Presence	0.697**	0.653**
Ease of Control	PENS Intuitive controls	0.856**	0.836**
Audiovisual appeal	AD Beauty	0.732**	0.815**
Challenge	PENS Competence	0.722**	0.325*
Goals and Rules	AD Pragmatic Quality	0.711**	0.733**
Progress feedback	AD Pragmatic Quality	0.643**	0.522**

A good discriminant validity showed that the factors were shown to be indeed sufficiently unrelated.

5.3.7. Self-Assessment Manikin (SAM) scale

The SAM is a tool to measure the emotional affective reaction in the dimensions pleasure, arousal, and dominance by letting the user point on pictures. [70] The Self-Assessment Manikin was used during balance exercises to measure the emotional state [69]. SAM was shown as a reliable and objective way to measure emotions in VR Gaming [71].

5.3.8. Game Experience Questionnaire (GEQ)

The Game Experience Questionnaire was developed 2007 by a European research project named FUGA [46]. It is widely used in at least 515 papers [57]. The structure was examined by using focus groups of infrequent and frequent gamers talking about their feelings on video games and experts building on these results and theoretical considerations. A questionnaire was created assessed by a 5-point Likert-scale from "not at all" (0) to "extremely" (4). Orignally ten factors and 92 items were reduced after the questionnaire was tested with 380 participants referring to a game of their choice. After exploratory factor analysis on the results, a 7-factor solution with 82 items was concluded. It explained 52 % of variance in all items with most questions havinge more than 30 % correlation to only one factor. The seven factors are:

- Immersion (previously two factors Sensory and Imaginary Immersion)
- Tension (said to unforseenly emerged from Negative Affect although the factor Suspense already existed)
- Competence (includes previous factor Experienced Control)
- Flow
- Negative Affect
- Positive affect
- Challenge

Two 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" .

A paper from 2018 did a systematic review using 73 studies that applied the GEQ [57]. 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 publicated. 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 condluded that the factor Negative Affect is not satysfying due to a low reliability coefficient and the factor Challenge just barely. A exploratory factor analysis was conducted and came to the conclusion to remove some items that did not load well on the given or any other factor. Renaming Flow to "loss of time" was proposed. 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 adress multiple issues. What are possible readons why the results of the conducted study differ from the original study? Are there maybe missing questions that could lead to 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?

Another validation study use exploratory and confirmatory factor analyses to find partial support for the GEQ and large support for the PENS [53], also proposing to summarize negative affect, tension/annoyance and challenge to one single factor negativity.

Another trivial problem of the GEQ is that there is another questionnaire in the field of game research with the same abbreviation GEQ. The Game Engagement Questionnaire also measures game experience, but rather the psychological engagement [58] using 19 items in four categories absorption, folow, presence and immersion. One item from the presence category is e.g. "Things seem to happen automatically". The development was supported by a test to distract participants playing a video game with a prerecorded statement and classifying their reaction. They found a hint on a negative relationship between the Game Engagement Questionnaire and the Dissociative Experiences Scale (DES) that measures dissociation eg that you find yourself in places without knowing how to got there [59].

5.3.9. Flow State Scale

The term flow was developed to understand the phenomenon of intrinsically motivated activity [60]. It was described as the ideal experience during work or play [61] In flow, a person is totally invovled in a task while challenge and skill level hold a perfect balance [62]. Jackson et al have further shown that, when experiencing flow, the challenge level was perceived and rated slightly higher with a value of 8.3 compared to the skill level with a value of 7.3 [63]. Getting in flow state would be pleasant for exergames too also because it plays a role in motivation and enjoyment and people seem to show peak performance while in flow state [63], [64]. To evaluate flow, there is a questionnaire called the "Flow State Scale". It originally consisted of 54 items with 9 factors but be reduced to 36 items while keeping a high reliabiliy with coefficient alphas rated above 0.80 [63] Here are the 9 original factors. A study with 1083 athletes proved an acceptable fit [65]. THe last two have the most potential for removal [63]:

1. Sense of Control
2. Challenge-Skill-Balance
3. Clear Goals
4. Merging with task
5. Unambigious feedback
6. Concentration
7. Loss of Self-Consciousness
8. Transformation of Time (weakest factor, also shown to not have acceptable internal consistency [66])
9. Intrinsic motivation (may be more of a higher-order factor)

Another study developed a flow state scale for occupational tasks in occupational therapy by letting the participants play video games and confirming that the computer game task used in the study represented measurable activities typically encountered in occupational therapy. By using a confirmatory factor analysis, they condluded three factors Sense of control, Positive emotional experience and Absorption by concentrating. The total score on their flow state scale was negatively correlated with the total score on the State-Trait Anxiety Inventory (STAI), indicating that the flow state was associated with lower anxiety levels.

5.4. Psychophysiological measures

5.4.1. 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 20" [67] "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" [67] "All these studies demonstrate that physiological signals are closely related to players' self-reported emotional states and behaviour, while Chanel 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" [67]

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" [67]. "A large number of studies have shown that psychophysiological measures can be used to index emotional, motivational and cognitive responses to media messages" [67] "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" [67] "Psychophysiological research is defined as using physiological signals to study psychological phenomena (Cacioppo et al. 2007: 5)." [67] 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 [68]. Also, these measurements can be taken continously during the gameplay and can be mapped to a recorded gameplay session afterwards [69] which allows great post-analysis. Nonetheless, they should not be relied on alone. The researcher should always ensure the correct interpretation of the signals [69], 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 [70]. 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 meausrement [69]. Also the demographic background of any person and the baseline of the measured variable in idle state should be taken into consideration [69] 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 [67].They can be more easily measured and concrete emotions can be derived from them [71] 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 excercise. 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

Liu et al [72] used physiological measured to estimate anxiety in order to adapt the difficulty of the played game in real time. Measures were: cardiovascular activity, including interbeat interval, relative pulse volume, pulse transit time, heart sound, and preejection period; electrodermal activity (tonic and phasic response from skin conductance) and EMG activity (from Corrugator Supercilii, Zygomaticus, and upper Trapezius muscles). Measured anxiety could be matched to subjective anxiety rating with 78 % accuracy. Also, the players received the game as more challenging and satysfiing as when the difficulty was changed based on the players performances. Additionally, their performance improved and the perceived anciety decreased.

5.4.2. Facial Electromyography (EMG)

Sensors are attached to the skin to measure the eletric and so physical activity of muscles. They can be used to detect facial expressions. It was shown that eye [73] and eyebrow muscles are a good indicator for positive and cheek muscles a good indicator for negative emotions [69], [74]. Hazlett conducted a study which showed, that during a car racing game, the zygomaticus muscle, which controls smiling, correlated well to positive events and corrugator muscle, which controls frowning, correlated well to negative events. [75] Disadvantages are discomfort wearing face sensors, the inability to talk during gameplay and unnatural facial expressions due to feeling the sensor [69] 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" [76]

5.4.3. 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 [69], [77]. The measuring device can function wireless and be made out of soft materials [78] do not be disturbing during exercise The main disadvante is some seconds of latency which makes it more difficult to link a signal to a cause [69]

"Significant covariation was obtained between (a) facial expression and affective valence judgments and (b) skin conductance magnitude and arousal ratings" [79] "most drivers studied, skin conductivity and heart rate metrics are most closely correlated with driver stress level" [66]

5.4.4. Cardiac activity / Heart Rate

Heart rate was shown to effectively recognise all four basic negative emotions anger, fear, disgust and sadness [80] and is tied to emotional arousal [69]. There are many possibilities to measure heart rates. Chest strap monitors are very precise and are affordable (75 €) [81] and also current smartwatches can deliver good results [82].

5.4.5. Heart Rate Variability

The variability of the heart rate can be also obtained from measuring the heart rate but needs further analysis [69]. There is emerging analysis for its role in regulated emotional responding [83] with higher variability typically correlating with better emotional regulation [source?]. However, it was shown that heart reate variability alone is not that effective in recognizing emotions with a baseline below 50 % [84].

5.4.6. Finger Pulse Volume (FPV)

FPV was shown to be sensitive to changed in experimentally creates anxiety [85].

5.4.7. Electroencephalography (EEG)

Electroencephalography is used to measure brain waves with a high temporal accuracy [69]. Some consider it as "ideal psychophysiological measure" as its "constantly fluctuating and continuously responsive to changing psychological stimuly" [26]. The activity of different brain regions can be visualized but there is a lot of interpreation possibilities, especially since the origin of the brainwaves is not apparent [69]. At least it can be used to measure a player's engagement in a video game, e.g. there was increased activity when diing in. Super Meat Boy [86]. With the Emotive EEG headset [87], 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)" [67] Alpha waves were found to be more prominent during tasks that did not require attention, such as mental arithmetic while beta waves were associated with emotionally positive or negative tasks, as well as cognitive tasks [88].

5.4.8. Respiration rate

Studies reported higher respiration rate during arousal [89], increases in inspiratory time when laughing and decreases during disgust [90]. Although this measure does not give as much insights as others, it can be well estimated from using heart rate sensors, so no addtional hardware is neccessary when heart rate is already measured. The root mean square error then is 0.648 per minute [91].

5.4.9. 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))[92]

5.4.10. 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)." [67]]

5.5. Comparing exergames with conventional interventions

5.5.1. Measures

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

A meta analyses with 17 studies showed similar anxiety level reduction compared to conventional interventions, but mostly no improvement as additional treatment to conventional interventions [93]. This evidence can be limited as 93% of studies did not report when the anxiety levels were assessed after the last exercise session. The anxiety was measured most often using the Hospital Anxiety and Depression Scale (HADS-A) followed by the State Anxiety Inventory (STAI) and Anxiety Inventory (BAI).

The same authors did a between-groups meta-analyses to examine the effect on muscle strength of exergames versus no intervention and usual care intervention [94]. This was defined as "conventional rehabilitations, home-based physical activity programs, resistance training, balance training, traditional multicomponent training, and other similar programs". Comparing exergame vs. no intervention showed no effects in handgrip strength in heathy/unhealthy middle--aged/older adults, knee extension maximum voluntary isometric contraction (MVIC) in healthy older adults. Exergames vs. usual care showed advantages in improvement of handgrip strength, knee flexion MVIC, elbow extension MVIC and propably upper and lower limb muscle strength, but no advantages for knee extension MVIC or elbow flexion MVIC. When exergaming was used additionally to usual care, it only provided improvement of handgrip strength in children with hemiplegic cerebral palsy.

This review systematically examined the role of exergames in reducing weight-related outcomes among obese children and adolescents [95]. 10 studies met the inclusion criteria. e.g. no studies with participants with other ongoing weight-losing interventions ang high quality regarding the tool for Quality Assessment of Controlled Intervention Studies designed by the National Heart, Lung, and Blood Institute's (NHLBI) [96]. Seven out of ten studies reporting better outcomes in the exergame than in the intervention groups. While there is a potential positive effect further research is needed to determine their effectiveness in childhood obesity treatment and identify the most effective approach. Weight-related outcomes were: weight/body mass index/z-scores, fat and/muscle mass and hip and waist circumference. Another review analysed the impact for childhood obesity in field-based studies [97] as naturalistic environments are advantageous in order to prove practical benefit [8]. It came to unclear results for both physical activity and weight-related outcomes due to issues regarding design, measurement and methodology. The results show that the examined studies often had small sample sizes and short durations or low frequencies per week, limiting the generalizability and practical implications of the findings. It is noteworthy that the inclusion of children and adolescents in all studies may be influenced by the availability of exergames designed for this age group and the rising concern over childhood obesity in the USA.

In another three month study with 117 college student [98], exergames lead to improvements in normalized diastolic blood pressure, sit-up tests and 3-min step tests but body fat percentage was increased. Regular exerciser showed push-ups improvement and irregular exercises improved response time.

"Being provided with the opportunity to engage in DDR or other exergames, children and adolescents became more genuinely interested and selfefficacious compared with conventional physical education classes such as fitness or aerobic dance "

"For example, it was found that EE values of DDR were equivalent to 7.0 metabolic equivalents (METs), which is in the moderate-to-vigorous intensity PA (MVPA?)" [97].

dual-energy X-ray absorptiometry [%fat and bone mineral density BMD] and magnetic resonance imaging. Cardiovascular risk factors included blood pressure, cholesterol, triglycerides, glucose and insulin. Exergaming reduced body fat and increased BMD among those adolescent girls [99].

In a 3-month-intervention study, several effects of exergaming on 29 older adults were examined [100]. In the outcomes measures, enduraced assessed by a 2 minutes step test increased significantly, moderate effects were measured in quality of life and lower body strength, small effects were detected in gait speed, mobility in the lower body and the balance capabilities.

Emotions: Anxiety

General: Weight/height/BMI

Muscle strength: isokinetic strength handgrip strength MVIC 1RM (4 from [94])

Behavior: Physical activiy (accelerometers)

MOTOR LEARNING Designing educational exergames should be based on the knowledge of motor control and motor learning mechanisms, considering motor learning as an internal process that evaluates individual ability and performance. Incorporating scientific paradigms of motor control and motor learning into exergame design is crucial for assessing the actual effectiveness of exergames in training and rehabilitation programs. [101]

GPT ***

Physical Health: **************** x Body Mass Index (BMI) x Body composition (e.g., percentage of body fat, lean mass) Cardiovascular fitness (e.g., VO2 max, heart rate) x Muscular strength (e.g., handgrip strength, leg strength) Muscular endurance Flexibility x Balance and coordination x Reaction time Agility Motor skills

Physical Activity and Energy Expenditure ** Total minutes engaged in physical activity Step counts Energy expenditure (calories burned) Metabolic Equivalent of Task (Met) (energy expandature relative to mass) Sedentary behavior (e.g., sitting time, screen time)

Cognitive Health and Function ****************************** Cognitive performance (e.g., attention, memory, executive function) Processing speed x Reaction time Task performance accuracy Problem-solving ability Visual-motor coordination Dual-task performance Balance: Berg Balance Scale Senior Fitness Test.

Psychological and Behavioral Measures ************************************ Enjoyment and engagement Motivation and self-efficacy Perceived exertion x Mood and affect Self-esteem Social interaction and cohesion Adherence and compliance to exercise programs Behavioral changes (e.g., physical activity levels outside of exergaming sessions)

Physical activity in inactive children

A study from 2010 has shown the big potential of exergames for a physical education classroom [13]. 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 although it is not stated which device was used for measuring. It was shown that the inactive children increased their physical activity from 1.6 min to 9.2 min during a lesson. This massive increase could be enabled by enhancing the time of opportunities for participating in excercises from 3.8 to 11.6 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 or it is realted to the novelties. 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 [5]. After six weeks of intervention, 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. [102]

PlayStation EyeToy intervention

Playing active video games can help overweight or obese children who were current users of sedentary video games to improve their body mass index. This improvement is most likely because their aerobic fitness gets better [103]. There were 162 participants in the intervention and 160 in the control group, all having a Playstation 2 or 3 gaming console but no fitness video games. They received the neccessary hardware and a selection of Sony PlayStation fitness video games that use EyeToy, an addtional camera for the console, to detect body movements. They were encouraged to engage in daily 60 minutes of exergaming in periods of inactivity or traditional gaming. The study went 24 weeks and after 12 weeks, a bunch of new games was sent to the intervention group to ensure sustainability. Outcome measurements were: BMI (by measuring height and weight) Body fat was assessed using standardized bio-electrical impedance analysis procedures Aerobic fitness was assessed using the 20 meter shuttle test [12,13]. Seven day physical activity was measured objectively using accelerometry [14] daily (for seven days) self-reported snack food consumption was assessed using a participant diary developed and tested in a previous pilot study [8]. Only aerobic fitness at 24 weeks met the conditions for mediation, and was a significant mediator of all treatment outcomes

Exergames vs standard exercise vs no exercise

One study [104] compared exercise videogames with standard exercise and control groups. The results showed that participants in the exercise videogame group had higher levels of moderate to vigorous physical activity and greater improvements in health risk indices, such as cholesterol and HbA1c, compared to the other groups. These findings suggest that exercise videogames can be an effective and enjoyable method to promote sustainable physical activity and achieve significant health benefits.

5.6. Evaluation of excersises

6. Evaluation methods for exergames

Exergames motivate the player to engage in physical activity, often by using modeling by a virutal trainer, feedback on the player's performance and reinforcement [105].

7. Conclusion

8. Prospect

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 adoloscents during social isolation; an overview”, 2020.

[2] A. Andrade, W. M. da Cruz, C. K. Correia, A. L. G. Santos, and G. G. Bevilacqua, “Effect of practice exergames on the mood states and self-esteem of elementary school boys and girls during physical education classes: A cluster-randomized controlled natural experiment”, 2020.

[3] Melchior, “Internet and video game use in relation to overweight in young adults”, 2014.

[4] Kowert, “Unpopular, overweight, and socially inept; reconsidering the stereotype of online gamers”, 2014.

[5] Lai, “Effects of interactive video-game based system exercise on the balance of the”, 2013.

[6] Khamzina, “Impact of pokemon go on physical activity”, 2019.

[7] `https://enwikipedia.org/wiki/List_of_most-played_mobile_games_by_player_count`, [Online; accessed 31-May-2023].

[8] T. Baranowski, D. Abdelsamad, J. Baranowski, T. M. O’Connor, D. Thompson, A. Barnett, E. Cerin, and T.-A. Chen, “Impact of an active video game on healthy children’s physical activity”, *Pediatrics*, vol. 129, no. 3, e636–e642, 2012.

[9] `https://www.re-mission2.org`, "[Online; accessed 31-May-2023]".

[10] Kato, “A video game improves behavioral outcomes in adolescents and young adults with cancer a randomized trial”, 2008.

[11] 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`.

[12] U. E. Bauer, P. A. Briss, R. A. Goodman, and B. A. Bowman, “Prevention of chronic disease in the 21st century: Elimination of the leading preventable causes of premature death and disability in the usa”, *The Lancet*, vol. 384, no. 9937, pp. 45–52, 2014.

[13] V. A. Fogel, R. G. Miltenberger, R. Graves, and S. Koehler, “The effects of exergaming on physical activity among inactive children in a physical education classroom”, *Journal of applied behavior analysis*, vol. 43, no. 4, pp. 591–600, 2010.

[14] S. R. Heaselgrave, J. Blacker, B. Smeuninx, J. McKendry, and L. Breen, “Dose-response relationship of weekly resistance-training volume and frequency on muscular adaptations in trained men”, *International Journal of Sports Physiology and Performance*, vol. 14, no. 3, pp. 360–368, 2019.

[15] Y. Oh and S. Yang, “Defining exergames & exergaming”, *Proceedings of meaningful play*, vol. 2010, pp. 21–23, 2010.

[16] C. J. Caspersen, K. E. Powell, and G. M. Christenson, “Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research.”, *Public health reports*, vol. 100, no. 2, p. 126, 1985.

[17] Peng, “Is playing exergames really exercising a meta-analysis of energy expenditure in active video games”, 2011.

[18] P. Caserman, “Full-body motion tracking in immersive virtual reality-full-body motion reconstruction and recognition for immersive multiplayer serious games”,

[19] S.-A. A. Jin, “Avatars mirroring the actual self versus projecting the ideal self: The effects of self-priming on interactivity and immersion in an exergame, wii fit”, *CyberPsychology & Behavior*, vol. 12, no. 6, pp. 761–765, 2009.

[20] Hortobagyi, “Effects of exercise dose and detraining duration of mobility at late midlife; a randomized clinical trial”, 2021.

[21] A. Drachen, “Behavioral telemetry in games user research”, *Game user experience evaluation*, pp. 135–165, 2015.

[22] H. R. Marston and P. A. McClenaghan, “Play yourself fit: Exercise+ videogames= exergames”, *Serious Games and Virtual Worlds in Education, Professional Development, and Healthcare*, pp. 241–257, 2013.

[23] J. Clay, M. Lockyer, P. Massey, and C. Fencott, *Game invaders: The theory and understanding of computer games*. John Wiley & Sons, 2012.

[24] A. Wols, A. Lichtwarck-Aschoff, E. A. Schoneveld, and I. Granic, “In-game play behaviours during an applied video game for anxiety prevention predict successful intervention outcomes”, *Journal of psychopathology and behavioral assessment*, vol. 40, pp. 655–668, 2018.

[25] J. M. Kivikangas, L. Nacke, and N. Ravaja, “Developing a triangulation system for digital game events, observational video, and psychophysiological data to study emotional responses to a virtual character”, *Entertainment Computing*, vol. 2, no. 1, pp. 11–16, 2011.

[26] B. Hatfield and D. M. Landers, “Psychophysiology in exercise and sport: An overview”, 1987.

[27] S. H. Boutcher and D. Landers, “Anxiety reduction in conditioned and unconditioned subjects following vigorous exercise”, 1986.

[28] H. DeVries and G. Adamas, “Electromyographic comparison of single doses of exercise and meprobamate as to effects on muscular relaxation”, 1972.

[29] D. Zillmann, R. C. Johnson, and K. D. Day, “Attribution of apparent arousal and proficiency of recovery from sympathetic activation affecting excitation transfer to aggressive behavior - sciencedirect”, 1974.

[30] S. Keller and P. Seraganian, “Physical fitness level and autonomic reactivity to psychosocial stress”, 1984.

[31] P. J. Teixeira, E. V. Carraça, D. Markland, M. N. Silva, and R. M. Ryan, “Exercise, physical activity, and self-determination theory; a systematic review.pdf”, 2012.

[32] E. Childs and H. de Wit, “Regular exercise is associated with emotional resilience to acute stress in healthy adults 05-00161”, 2015.

[33] Teixeira and A. L. Palmeira, “Needs satisfaction effect on exercise emotional response; a serial mediation analysis with motivational regulations and exercise intensit”, 2016.

[34] J. H. Kerr and G. Kuk, “The effects of low and high intensity exercise on emotions, stress and effort”, 2000.

[35] J. Li, Y.-L. Theng, S. Foo, and X. Xu, “Exergames vs. traditional exercise investigating the influencing mechanism of platform effect on subthreshold depression among older adults”, 2017.

[36] Makivic, “Heart rate variability (hrv) as a tool for diagnostic and monitoring performance in sport and physical activities”, 2013.

[37] 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.

[38] Reeves, “Endogenous hyperthermnia in normal human subjects; i. experimental study of evoked potentials and reaction time”, 1985.

[39] H. K. Hammond and V. F. Froelicher, “Normal and abnormal heart rate responses to exercise”, 1985.

[40] Haskell, “Updated recommendation for adults from the american college of sport midicine and the american heart association”, 2007.

[41] F. C. Blumberg, S. F. Rosenthal, and J. D. Randall, “Impasse-driven learning in the context of video games”, *Computers in Human Behavior*, vol. 24, no. 4, pp. 1530–1541, 2008.

[42] H. Desurvire and C. Wiberg, “Game usability heuristics (play) for evaluating and designing better games: The next iteration”, in *Online Communities and Social Computing: Third International Conference, OCSC 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings 3*, Springer, 2009, pp. 557–566.

[43] H. Desurvire, M. Caplan, and J. A. Toth, “Using heuristics to evaluate the playability of games”, in *CHI’04 extended abstracts on Human factors in computing systems*, 2004, pp. 1509–1512.

[44] B. Strååt and H. Warpefelt, “Applying the two-factor-theory to the play heuristics.”, in *DiGRA Conference*, 2015.

[45] M. Alshmemri, L. Shahwan-Akl, and P. Maude, “Herzberg’s two-factor theory”, *Life Science Journal*, vol. 14, no. 5, pp. 12–16, 2017.

[46] Poels, “Game experience questionnaire”, 2007.

[47] Rigby, “The player experience of need satisfaction (pens)”, 2007.

[48] 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 childrent”, 2020.

[49] 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.

[50] 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.

[51] E. L. Deci and R. M. Ryan, “Intrinsic motivation and self-determination in human behavior”, 1985.

[52] B. Wienke and D. Jekauc, “A qualitative analysis of emotional facilitators in exercise”, 2016.

[53] 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.

[54] M. E and M. D., “Variations in self-determination across the stages of change for exercise in adults. motivation and emotion”, 1997.

[55] V. V. Abeele, K. Spiel, L. Nacke, D. Johnson, and K. Gerling, “Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences”, *International Journal of Human-Computer Studies*, vol. 135, p. 102 370, 2020.

[56] M. Hassenzahl, M. Burmester, and F. Koller, “Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität”, *Mensch & Computer 2003: Interaktion in Bewegung*, pp. 187–196, 2003.

[57] 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.

[58] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny, “The development of the game engagement questionnaire: A measure of engagement in video game-playing”, *Journal of experimental social psychology*, vol. 45, no. 4, pp. 624–634, 2009.

[59] D. B. Wright and E. F. Loftus, “Measuring dissociation: Comparison of alternative forms of the dissociative experiences scale”, *American Journal of Psychology*, vol. 112, no. 4, pp. 497–519, 1999.

[60] J. Nakamura and M. Csikszentmihalyi, “Handbook of positive psychology chapter 7; the concept of flow”, 2002.

[61] M. Csikszentmihalyi and I. S. Csikszentmihalyi, *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press, 1992.

[62] Jackson, “Psychological correlates of flow in sport”, 1998.

[63] Jackson, “Development and validation of a scale to measure optimal experience the flow state scale”, 1996.

[64] D. J. Harris1, K. L. Allen, S. J. Vine, and M. R. Wilson, “91 a systematic review and meta-analysis of the relationship between flow states and performance”, 2020.

[65] N. A. Stavrou and Y. Zervas, “Confirmatory factor analysis of the flow state scale in sports”, 2011.

[66] S. P. VLACHOPOULOS, C. I. KARAGEORGHIS, and P. C. TERRY, “Confirmatory factor analysis of the flow state scale in sports”, 2000.

[67] J. M. Kivikangas, G. Chanel, B. Cowley, I. Ekman, M. Salminen, S. Järvelä, and N. Ravaja, “A review of the use of psychophysiological methods in game research”, *journal of gaming & virtual worlds*, vol. 3, no. 3, pp. 181–199, 2011.

[68] Hamdi, “Emotion recognition based on heart rate and skin conductance v6”, 2015.

[69] Nacke, “Physiological game evaluation”, 2015.

[70] Javorka, “Heart rate recovery after exercise”, 2002.

[71] Seo, “Automatic emotion-based music classification for supporting intelligent iot applications”, 2019.

[72] C. Liu, P. Agrawal, N. Sarkar, and S. Chen, “Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback”, *International Journal of Human-Computer Interaction*, vol. 25, no. 6, pp. 506–529, 2009.

[73] Ravaja, “Phasic emotional reactions to video game events”, 2018.

[74] Mandryk, “Using psychophysiological techniques to measure entertainment”, 2006.

[75] R. L. Hazlett, “Measuring emotional valence during interactive experiences: Boys at video game play”, in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2006, pp. 1023–1026.

[76] M. Perusquía-Hernández, S. Ayabe-Kanamura, K. Suzuki, and S. Kumano, “The invisible potential of facial electromyography”, 2019.

[77] M. E. Dawson, A. M. Schell, and D. L. Filion, “The electrodermal system.”, 2017.

[78] H. Kim, S. Kwon, Y.-T. Kwon, and W.-H. Yeo, “Soft wireless bioelectronics and differential electrodermal activity for home sleep monitoring”, *Sensors*, vol. 21, no. 2, p. 354, 2021.

[79] P. J. LANG, M. K. C.REENWALD, M. M. BRADLEY, and A. O. HAMM, “Looking at pictures: Affective facial visceral and behavioral reactions”, 1993.

[80] R. W. Levenson, “Blood, sweat, and fears: The autonomic architecture of emotion”, *Annals of the New York Academy of Sciences*, vol. 1000, no. 1, pp. 348–366, 2003.

[81] https://www.amazon.de/Polar-Herzfrequenz-Sensor-Bluetooth-Wasserdichter-Brustgurt/dp/B07PM54P4N, "[Online; accessed 31-May-2023]".

[82] Gillinov, “Variable accuracy of wearable heart rate monitors during aerobic exercise”, 2017.

[83] Appelhans, “Heat rate variability as index of regulated emotional responding”, 2006.

[84] Ferdinando, “Emotion recognition by heart rate variability”, 2014.

[85] L. J. Bloom and G. M. Trautt, “Finger pulse volume as a measure of anxiety; further evaluation”, 1977.

[86] McMahan, “Evaluating electroencephalography engagement indices during video game play”, 2018.

[87] https://www.emotiv.com/epoc-flex/, [Online; accessed 31-May-2023].

[88] W. J. Ray and H. W. Cole, “Eeg alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes”, 1985.

[89] Homma, “Breathing rhythms and emotions”, 2008.

[90] Boiten, “The effects of emotional behaviour on components of the respiratory cycle”, 1998.

[91] Natarajan, “Measurement of respiratory rate using wearable devices and applications to covid-19 detection”, 2021.

[92] Y. Lu, W.-L. Zheng, B. Li, and B.-L. Lu, “Combining eye movements and eeg to enhance emotion recognition”, 2015.

[93] R. B. Viana, S. J. Dankel, J. P. Loenneke, P. Gentil, C. A. Vieira, M. d. S. Andrade, R. L. Vancini, and C. A. B. de Lira, “The effects of exergames on anxiety levels: A systematic review and meta-analysis”, *Scandinavian Journal of Medicine & Science in Sports*, vol. 30, no. 7, pp. 1100–1116, 2020.

[94] R. B. Viana, V. N. de Oliveira, S. J. Dankel, J. P. Loenneke, T. Abe, W. F. da Silva, N. S. Morais, R. L. Vancini, M. S. Andrade, and C. A. B. de Lira, “The effects of exergames on muscle strength: A systematic review and meta-analysis”, *Scandinavian Journal of Medicine & Science in Sports*, vol. 31, no. 8, pp. 1592–1611, 2021.

[95] F. Valeriani, C. Protano, D. Marotta, G. Liguori, V. Romano Spica, G. Valerio, M. Vitali, and F. Gallè, “Exergames in childhood obesity treatment: A systematic review”, *International Journal of Environmental Research and Public Health*, vol. 18, no. 9, p. 4938, 2021.

[96] *National heart, lung, and blood institute’s. study quality assessment tools*, https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools, [Online; accessed 12 Juni 2023].

[97] Z. Gao and S. Chen, “Are field-based exergames useful in preventing childhood obesity? a systematic review”, *Obesity Reviews*, vol. 15, no. 8, pp. 676–691, 2014.

[98] H.-C. Huang, M.-K. Wong, J. Lu, W.-F. Huang, and C.-I. Teng, “Can using exergames improve physical fitness? a 12-week randomized controlled trial”, *Computers in Human Behavior*, vol. 70, pp. 310–316, 2017.

[99] A. E. Staiano, A. M. Marker, R. A. Beyl, D. S. Hsia, P. T. Katzmarzyk, and R. Newton, “A randomized controlled trial of dance exergaming for exercise training in overweight and obese adolescent girls”, *Pediatric obesity*, vol. 12, no. 2, pp. 120–128, 2017.

[100] S. Neumann, U. Meidert, R. Barberà-Guillem, R. Poveda-Puente, and H. Becker, “Effects of an exergame software for older adults on fitness, activities of daily living performance, and quality of life”, *Games for Health Journal*, vol. 7, no. 5, pp. 341–346, 2018.

[101] P. A. Di Tore and G. Raiola, “Exergames in motor skill learning”, *Journal of physical education and sport*, vol. 12, no. 3, p. 358, 2012.

[102] Hardy, “Adaption model for indoor exergames”, 2011.

[103] R. Maddison, A. J. Cliona Ni Mhurchu and, H. Prapavessis, L. S. Foley, and Y. Jiang, “Active video games the mediating effect of aerobic fitness on body composition”, 2012.

[104] B. C. Bock, S. I. Dunsiger, J. T. Ciccolo, E. R. Serber, W.-C. Wu, P. Tilkemeier, K. A. Walaska, and B. H. Marcus, “Exercise videogames, physical activity, and health: Wii heart fitness: A randomized clinical trial”, *American journal of preventive medicine*, vol. 56, no. 4, pp. 501–511, 2019.

[105] C. H. Elizabeth Jane Lyons, “Prevalence of behavior changing strategies in fitness video games: Theory-based content analysis”, 2013.

A. Appendix

The appendix goes here... english