Master Thesis in
Human Computer Interaction

Balancing Multiplayer Exergames in Virtual Reality with Dynamic Difficulty Adaptation

Tamara Mantz
Student ID: 1122584

Supervisors
First supervisor: Dr. H.J. (Hanna) Hauptmann
Second supervisor: Dr. W.O. (Wolfgang) Hürst

Utrecht University

Graduate School of Natural Sciences
March 2024

Abstract

Regular physical activity has many health benefits, including preventing cardiovascular disease and other conditions linked to obesity. However, people often lose motivation to incorporate sports into their lives over time. Exergames, which combine physical activity with games, have the potential to motivate people to do physical activities. In particular, multiplayer games can be motivating due to their competitive nature and by connecting friends remotely.

However, one issue remains: differing skills among players can lead to frustration or boredom, particularly when skill levels vary significantly. As a solution, this thesis proposes dynamically adapting the difficulty of the game during game play for players individually.

Previous research on balancing exergames has identified negative effects, such as overbalancing or a detrimental impact on the self-efficacy of weaker players. To address these issues and develop an approach for dynamic difficulty adaptation aligned with players' needs, the research question for this thesis is formulated as follows: How can dynamic difficulty adaptation be used to balance mismatched skills among multiple players? To achieve this, it was essential to engage users in the design process, which led to several ideas created according to requirements for dynamic difficulty adaptation, which have been derived from preliminary focus groups. A prototype was then developed which incorporated dynamic difficulty adaptation in a game similar to the Virtual Reality game Beat Saber.

The second part of the thesis involved an empirical cycle to investigate effects on perceived fairness, competence, and motivation. User experiments were conducted with 14 participant pairs playing against each other in a within-subjects study. Results of the quantitative analysis revealed that fairness was significantly higher for the game with dynamic difficulty adaptation, confirming one of the hypotheses. However, no differences were found for competence and motivation. Qualitative analysis of the interviews showed that participants felt motivated during the game with the adaptation approach, mainly because they could change the game's results at any point with more effort. Additionally, no overbalancing occurred and participants did not feel less competent due to receiving help during the game. However, some participants still prefer to compete in games without difficulty adaptation to compare true skills. The aforementioned design requirements were compared with the user study results and have mostly led to positive results. However, it is unclear if the same requirements hold true in other types of exergames. Particularly, focusing on creating close games and letting players control the received assistance led to the positive result of higher fairness. Future research can focus on applying these requirements in other contexts and examining long-term effects on motivation when using dyanmic difficulty adaptation.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

Physical activity is an essential component of a healthy lifestyle, serving as a preventive measure against obesity [63] and cardiovascular disease [34]. Furthermore, it is recommended to be moderately physically active for at least 150 minutes or engage in intense exercise for 75 minutes per week [21]. To counteract the negative implications of sedentary lifestyles, it is important to motivate individuals for the long term to engage in physical activities. However, it has been observed that up to 65% of individuals who start to exercise are prone to dropping out [7, 8, 29]. A common reason for discontinuation is time constraints [32], a challenge that could be mitigated through short, home-based exercises as there is no need to commute to another place such as a gym. Additionally, maintaining motivation can be problematic. Exercise games (exergames) have the potential to solve both problems by combining elements of computer games with physical activity, where movement can be tracked and used to control the game. Individuals might exercise more often due to the enjoyment of the game and can do so from home.

Exergames are now commercially available on consoles equipped with motion sensors in controllers and Virtual Reality (VR) head-mounted devices. VR offers a wide range of exergames, such as Beat Saber, FitXR, and Supernatural, as well as exergames that simulate real sports, such as table tennis or boxing. Rhythm exergames focus on striking and evading targets by moving arms and the upper body synchronized to music. Furthermore, multiplayer exergames have shown the potential to increase motivation. In competitive exergames, players often significantly increase their physical effort in order to win [61, 76, 49, 6, 75]. Additionally, exergames allow players to compete against their friends even while being apart.

However, competing with friends with different skills and physical abilities can lead to unbalanced games, where the less skilled player continuously loses. Thus, a mismatch of skills between players may lead to frustration of less skilled players or boredom of skilled players. Previous research has also shown that competitive computer games are more enjoyable for players if the game has been close and the outcome of the game is unclear until the end [27, 87].

As a solution, exergames can be modified according to players' abilities and skills in order to assist players who are less skilled. Therefore, frustration or early dropouts of regular exercises using exergames can be avoided. More skilled players benefit from a challenging game as boredom can be avoided while being more encouraged to make an effort. Thus, balancing exergames can also be achieved by increasing the difficulty of the game for more skilled players. It has been observed that a noticeable difficulty increase has a positive effect on players. Skilled players feel recognized and rewarded for good performance [41].

Previous research has focused on adjusting game difficulty, such as using players' heart rates as input for effort [77], increasing scores, changing target distance and size, or decreasing the number of targets [3, 35, 41]. However, one disadvantage of adapting difficulties is that they often become apparent to players. Therefore, adapting difficulties can decrease their self-efficacy in the case of less skilled players and can be perceived as unfair in the case of skilled players [35, 41].

Research that focused on balancing exergames primarily investigated effects within non-VR games, where players have to be physically together to play. Thus, players can see their fellow players moving. Competitive physical activities may cause individuals to feel observed and worried about negative reactions from peers [62]. It has already been observed that exergames reduce the barrier for physical activities in groups, as individuals are more focused on the screen [79, 67, 12]. This effect might be further enhanced when playing VR games, as it is not possible to see how other players are moving in real life. It might be possible that players feel less inclined to compare themselves with other players. Thus, it is unclear if difficulty balancing has the same negative effects as observed in previous research.

1.2 Problem Statement

Ultimately, the goal is to motivate individuals to be more physically active. Exergames, when played together with friends, have the potential to engage individuals long-term. Its simplicity makes it easy to learn and play for a wide audience. It has been also shown that motivating individuals to be physically active through exergames can be achieved better by simple games [74]. However, when players of different skill levels play together, it can lead to frustration or boredom for skilled and less skilled players. Thus, adapting the difficulty during the game based on the performance of each player, might mitigate these negative effects. The goal of the game is to keep each player challenged enough to have an exciting match but without feeling overwhelmed when playing with a more skilled opponent.

Adaptation of difficulty in exergames might have positive effects on players, which can lead to a higher probability that individuals will keep playing long-term. To verify that difficulty adaptation has the desired effects, this thesis focuses on investigating what effects adapting game difficulty has on players in VR exergames when playing in multiplayer mode.

Difficulty adapatation can be implemented in various ways and has to be tailored to the context of the game: players' personalities, their preference for difficulty adaptations and the particular game. A badly designed adaptation can lead to rejection of the exergame [4]. In this thesis rhythm exergames with difficulty adaptation will be investigated. An example of such a game is Beat Saber, where players hit cubes with a light saber in the rhythm of a song. By doing so players swing arms, move their upper body, crouch down and make steps to the side. To successfully play these exergames, general fitness as well as coordination is needed. Rhythm exergames allow for comparabiltiy between players as they have to do the same moves when competing with each other. As first part of this thesis, it is necessary to investigate what requirements should be taken into account when designing difficulty adaptation for rhythm exergames in VR.

The requirements will then be used to derive a design for difficulty adaptation that players are more likely to accept. To evaluate that the implemented solution has the desired positive effects on players, a user study has to be conducted. Therefore, this study consists of two parts. The first part is concerned with deriving requirements for designing as well as implementing a playable solution for the problem of mismatching skills in multiplayer mode. The second part is concerned about evaluating the solution and investigating the effects.

Apart from one study conducted by Jensen & Grønbæk [41], previous research has mostly focused on static difficulty adaptation. Players have received assistance or a handicap that did not adapt dynamically according to player's current performance during gameplay. These methods have led to overbalancing in some of the trials and were therefore not considered fair by participants [3]. Stach et al.'s study has focused on rewarding players based on increased heart-rates but suffered from overbalancing as well [77]. To mitigate overbalancing this research will focus on dynamic adaption of exergames during gameplay.

Another positive aspect of exergames is the player's focus on the screen rather than other players. While individuals can feel intimidated to engage in physical activities in front of peers, this negative effect has been less observed in previous research concerning exergames [62]. Thus, this particular issue might be even less present in VR as it is not possible to observe others' movements in the real world. Therefore, it is hypothesised that these negative effects will not occur and, instead, difficulty balancing will have positive effects on players' self-efficacy. Competence will be measured to investigate if this effect is present. VR exergames have not been previously researched when it comes to balancing exergames and may lead to different results due to its immersive characteristics.

As the goal of difficulty adaptation is to create a close game between players, skilled players might perceive it as unfair since they expect to perform better if they acquired higher skills through practice or a more physically active lifestyle. In similar research before, fairness has been measured as one of the dependent variables due to the possibility that players become demotivated when exergames provide unfair conditions [35, 41]. Difficulty adaptation can lead to the opposite of the desired effect if it is perceived unfair. Thus, perceived fairness has to be measured in order to validate if difficulty adaptation in VR exergames have a positive effect on players' motivation.

The primary aim of this research is to determine, what effects dynamic difficulty adaptation in VR exergames have to provide insights into how exergames in VR should incorporate difficulty adaptation to continuously motivate individuals to play exergames. For this, effects on motivation will be measured as well. This leads to three dependent variables that are relevant to determine positive effects in difficulty adaptation: motivation, fairness and competence. Therefore, this thesis contributes to current research by defining design requirements for the specified scope of rhythm exergames in VR and investigating the effects of the resulting solution. If positive effects on players can be confirmed, the derived requirements in this thesis can be used for designing fair and motivating exergames with dynamic difficulty adaptation.

1.3 Research Question

Among possible solutions for the problem, dynamic difficulty adaptation was chosen. It has the potential to keep players in a highly engaged state (flow state [25]) regardless of the individual skill levels of competing players. In order to balance exergames for mismatching skills of players, the effects of this solution will be measured. Thus, the research question is: *How can dynamic difficulty adaptation be used to balance mismatching skills of multiple players?*. To answer this research question, three areas have to be investigated: Competence, Motivation and Fairness.

Overall, a good player experience is needed in order for individuals to continue playing exergames long-term, which is a result of an intrinsically motivating game. Self-determination theory (SDT) describes what factors lead to intrinsic motivation, which is the core motivation behind sports and games [69]. Among other factors such as autonomy, perceived competence is crucial to feel intrinsically motivated. Player experience can be measured with scales such as the Player Experience of Need Satisfaction (PENS), which is based on SDT, and the Player Experience Inventory (PXI). PENS includes a competence construct as well as a game enjoyment construct [69]. It was adapted from the Intrinsic Motivation Inventory (IMI), which overall measures intrinsic motivation by taking into account perceived competence and interest/enjoyment [52]. Similarly, PXI includes constructs for challenge and mastery, which can be mapped to competence [1]. Other constructs of player experience questionnaires include immersion(PXI)/presence(PENS, ease of control(PXI))/ease of use(PENS), autonomy(PXI and PENS), audiovisual appeal (PXI) and more. These specific constructs are likely dependent on the chosen game's mechanics and aesthetics. Immersion/presence is likely to be high due to the immersive nature of Virtual Reality head-mounted displays [18, 84]. Therefore, it is expected that motivation and competence will differ when enhancing an exergame with dynamic difficulty adaptation. Fairness is added to measure if dynamic adaptation created overbalancing problems.

1.3.1 Competence

If the game is not adapted and the opponent is more skilled than the player, game difficulty also increases. Winning becomes difficult with a very skilled opponent. However, by adjusting the difficulty in-game this negative effect could be mitigated. Thus, a positive outcome of dynamic difficulty adaptation could be that players feel more competent since the difficulty is constantly adjusted according to their performance. Players can always reach goals in the game by putting their skills into use. Therefore, it may increase self-efficacy. Individuals with higher perceived competence are likely to have higher intrinsic motivation when doing sports. They expect as well as seek rewarding outcomes in sports [89]. This can lead to a higher likelihood that players with high perceived competence will continue playing exergames. Therefore, it will be investigated if dynamic difficulty adaptation has a positive effect on perceived competence. This leads to the sub-question R1: *Does dynamic difficulty adaptation increase perceived competence in players?*.

Hypothesis H_0: (H1.0) Dynamic difficulty adaptation in VR multiplayer exergames does not have an effect on players' perceived competence during gameplay.

Hypothesis H_1: (H1.1) Dynamic difficulty adaptation in VR multiplayer exergames increases players' perceived competence during gameplay.

1.3.2 Motivation

A motivating exergame is necessary in order to keep individuals engaged during gameplay and to increase the likelihood that the exergame will be repeatedly played. Players may increase their physical effort if they feel engaged in the game as they experience enjoyment and are motivated to win the game. Thus, players might be more physical active and have a higher heart-rate, which increases cardiovascular fitness [20]. For this, it is necessary to answer the sub-question R2: *Does dynamic difficulty adaptation motivate players?*.

Hypothesis H_0: (H2.0) Dynamic difficulty adaptation in VR multiplayer exergames does not have an effect on players' motivation.

Hypothesis H_1: (H2.1) Dynamic difficulty adaptation in VR multiplayer exergames increases players' motivation.

1.3.3 Fairnesss

Fairness is particularly important for difficulty adaptation as it adapts the game for each player differently. It might be the case that a player perceives it as unfair when another player receives an advantage. Previous research has observed that players may feel unfairly treated when the game's difficulty is adapted per player [3]. This was the case, when the game was overbalanced, e.g. making it easy for the less skilled player and difficult for the skilled player to win. Unfair games create a negative response and thus, are likely to not be played anymore. Therefore, the proposed solution is an adaptation that is implemented for every player and adapts to each player differently depending on their performance and the difference between the players' skills. It is therefore important to measure the perception of fairness when dynamic difficulty adaptation is present. The third sub-question R3 is: *Is dynamic difficulty adaptation perceived fair in players?*.

Hypothesis H_0: (H3.0) Dynamic difficulty adaptation in VR multiplayer exergames does not have an effect on players' perception of fairness of the game.

Hypothesis H_1: (H3.1) Dynamic difficulty adaptation in VR multiplayer exergames increases the players' perception of fairness of the game.

2 Literature Review

This reviewed literature gives an understanding of concepts related to the research topic and explains how its scope was defined. The scope entails competitive multiplayer exergames in virtual reality and how to balance games. The literature review was conducted by firstly defining the areas relevant for this research to be reviewed. For each area various search terms were used, which can be seen in Table 1. Literature was found through a Google Scholar search. Relevant literature found would then be examined for references. Thus, snowballing led to the review of further literature.

Area	**Search Terms**
Serious games	serious games definition, serious games motivation effect, gamification, e-learning, game-based learning, digital game-based learning
Exergames	exergames definition, multiplayer exergames, exergames competition, exergames motivation effect, balancing exergames
Virtual Reality	virtual reality definition, virtual reality exergames, virtual reality multiplayer exergames
Questionnaires	questionnaire motivation exergames, physical activity questionnaire, questionnaire self-esteem evaluation

Table 1: Search terms used

The literature review starts with defining what serious games are and how they can be used to motivate. A subcategory of serious games is exergames, which this thesis focuses on. Therefore, a definition of exergames follows. The second part explains the benefits of multiplayer exergames and how it affects motivation. Afterwards, the results of prior research is summarised. The summary includes balancing multiplayer exergames and its positive and negative effects on players. Lastly, this review will cover an explanation why VR seems to be a promising technology for exergames and how it can encourage players to stay active with exergames.

2.1 Serious Games

Serious games can be defined as games which serve another purpose other than entertainment. While definitions vary between authors, their common ground is that serious games entail a game that consists of an imaginary world, which might have relations to the real world and catches player's attention [53]. By doing this, serious games offer the possibility to persuade, educate or motivate players for a certain topic. Serious games are used in many areas: education, military, corporal training, health and others. The market of serous games was valued in 2020 to be $5.94 billion and has been estimated to grow significantly due to the Covid-19 Pandemic as educational institutes had to shift to online learning [88]. This further opened the path for digital tools in education and thus, for educational serious games.

Abt has coined the concept for the first time [2] and according to Abt, even games that were developed for pure entertainment can be used with a serious purpose in mind. Abt used board games, sports and early computer simulations in order to educate [30]. Therefore, serious games are not necessarily tied to digital games, nor do games have to be specifically developed for a serious purpose. Games have been utilized for training even before Abt provided a definition for the term "Serious Games". Wargames in World War II were put to practise in order to improve the U.S. Army's reputation within the population [45]. Since Sawyer connected it to digital games in 2002 the term is now mostly associated with digital games [70].

Michael and Chen provided a definition stating that serious games serve education as primary goal rather than entertainment [53]. However, education seems limiting as serious games can be also used for causes other than improving or learning skills. An example would be distracting patients from pain while playing a game [19]. Susi et al. defined serious games as games aiming to achieve another purpose other than entertainment while the player may or may not be aware of its purpose [81]. According to this definition, games with the goal to spread propaganda can be seen as serious games as well [83]. Lamaarti et al.'s definition includes three parts consisting of multimedia, entertainment and experience. It describes a serious game as a medium that entails multimodal interaction in order to create an experience for the player [45]. According to Alvarez and Djaouti serious games entail gameplay and a purpose other than entertainment as well as a sector. The sector describes the market in which the serious game is used [5].

Similar or overlapping domains to serious gaming exist such as *e-learning*, which refers to computer-supported learning usually in the form of interactive software [44, 54]. Therefore, serious games with the aim to educate can be also categorised as e-learning systems. However, e-learning does not necessarily have to include games. It may also consist of an application where digital educational content is shared or an application that facilitates holding online lectures.

Another related domain is *gamification*, which brings game elements into applications that are not designed as a game. According to Deterding et al., gamification can be described as "the use of game design elements in non-game contexts" [28]. Such elements can be leader boards and rewards. Gamification only enhances applications but does not make an application a (serious) game.

Game-based learning (GBL) is a similar domain to serious games as it focuses on using games for education. Thus, GBL can be seen as a domain within serious games as the latter may also entail other purposes than education. Some authors see no distinction between GBL and serious games [81, 23]. Similar to GBL is *digital game-based learning*, which is GBL but limited to digital games [81]. However, it emphasizes the aspect that students have become "digital natives" and therefore experienced in using digital media [66]. Using digital games for motivation purposes comes therefore natural to individuals, who can be described as "digital natives".

2.2 Exergames

Sedentary lifestyles have been proven to cause health issues. Office workers are prone to sedentary lifestyles as they spend the vast majority of the day at a desk. Overweight and the resulting cardiovascular diseases are known consequences [63, 34]. Exercise games (exergames) are serious games which focus on encouraging physical activity through gaming and therefore mitigate the negative consequences of a sedentary lifestyle. Motion sensors in gaming controllers are now widely commercially available and were made popular with Nintendo's release of the motion-controlled console Wii in 2006 [58]. Players are able to control games by using body movements which can consist of movements of the upper body and arms, jumps, running/walking on place or making steps to the side, front or back. Players can easily play exergames in their own home since exergames do not require more space than a home workout training. Therefore, exergames provide a way of exercising without the need to go to a gym. Moreover, online multiplayer modes make it possible to play with multiple individuals such as friends and family which adds competition as another motivating factor [37].

Motion control should feel as natural as possible to immerse players in the virtual world and increase their perceived presence within the game [46]. As physical exhaustion can occur, Sinclair et al. suggested that exergames should be kept simple and be designed with focus on the motion control itself. An example of this would be Dance Dance Revolution which is a rhythm exergame that includes a dance mat. In this rhythm dance game players have to step on the arrows shown on the mat in the right sequence. Sinclair et al. mentioned that avoiding physical exhaustion can also be achieved by adapting a game to players skills and fatigue during gameplay [74].

While it has been shown in previous research that exergames include light to moderate physical activity, it could be also seen that it increases visuospatial skills such as coordination and spatial awareness, motor skills, cognitive skills and even social behavior in case of multiplayer exergames [38, 72]. In order for exergames to achieve positive effects, it has to include motivational affordances that will in turn lead to the desired psychological outcomes such as achievement and self-esteem. According to Matallaoui et al., these affordances are game elements such as competitive multiplayer, leader boards or badges [51]. An exergame that includes multiple of such game elements is likely to lead to positive results and therefore can motivate players long-term. Thus, the novelty effect, which describes the phenomenon that players lose interest in exergames after a short amount of time, can be overcome [51]. A common criticism of serious games promoting physical activities is that they are designed to incorporate extrinsic motivation. Examples of such extrinsic motivators are the aforementioned leader boards and badges. Additionally, exergames should be designed similarly to games that were created purely for entertainment to motivate players for long-term play [60].

2.2.1 Flow Theory

Flow theory is particulary important for designing motivating exergames as it describes the rationale for being fully engaged in an activity. Focused concentration, the perception of time passing faster, joint action and awareness, as well as losing awareness of oneself as a social actor, are characteristics of being in a state of flow [56]. The complete list of conditions to enter a flow state can be seen in Table 2. The ultimate goal of the activity becomes less important or even irrelevant, as the process of reaching the end goal is inherently rewarding. Thus, activities do not necessarily require a rewarding outcome.

Condition
The activity contains a task that can be completed with the current capabilities
Being able to concentrate on given task
The task has clear goals that are reachable along the process of the activity
Instant feedback is provided during the activity
Feeling of being in control over one's actions
Feeling fully involved in the task
Being less aware of oneself during the task
Time is perceived faster

Table 2: Conditions for the Flow state [82]

An example of this phenomenon would be a player fully absorbed into a game world and playing for the enjoyment of the game without any external rewards. To enter a state of flow, it is important to be challenged without feeling overwhelmed and to receive immediate feedback, as well as having short goals that can be achieved during the activity. Flow depends on the balance between a person's capabilities and the challenges faced during the activity. In Figure 1, it can be seen that flow is a state outside of boredom and anxiety, given that capabilities and challenges are balanced. On one hand, capabilities exceeding challenges by far will likely lead to boredom. On the other hand, challenges being too difficult for the present capabilities will likely induce anxiety.

As the activity is intrinsically rewarding while being in flow, activities inducing flow are sought to be replicated. Sports and games make flow states more likely as they provide frequent and instant feedback and the possibility to achieve smaller goals along the process [56]. Thus, exergames that take into account conditions that likely induce flow can motivate individuals to play regularly. For this, the balance between challenges and skills is particularly important. In multiplayer games, another level is added to the individual skills of a player and the challenges a game has to offer: the differences between skills of multiple players. Thus, challenges can increase or decrease depending on the opponents' skills.

Additionally, exergames need to balance players' fitness and the provided exercises within the game. Thus, exergames have to be carefully designed in order to be motivating and not create fatigue during gameplay [74].

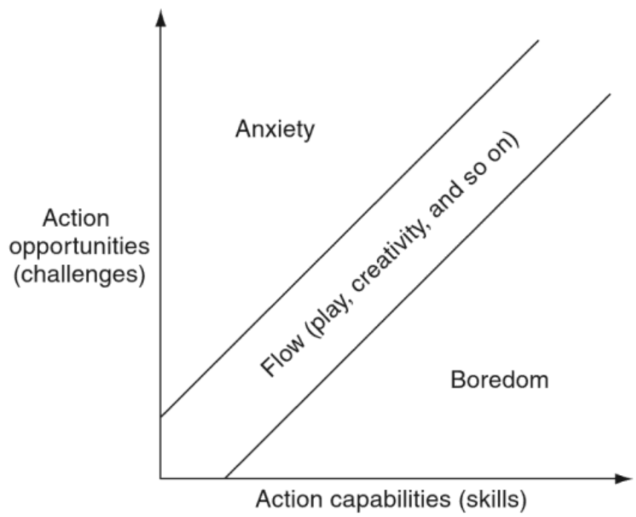


Figure 1: Balance between anxiety and boredom in flow [56]

2.2.2 Self-Determination Theory

Previous research has argued that intrinsic motivation underlies the interest in playing computer games and sports [69]. Thus, when players experience intrinsic motivation, they engage in games without expecting external rewards. The game experience becomes rewarding enough to sustain their interest in playing. SDT explains the factors that influence intrinsic and extrinsic motivation. Within SDT, the *basic psychological need theory* outlines three needs essential for fostering intrinsic motivation and contributing to overall well-being [69].

One of these needs is *autonomy*, which describes the need to willingly engage in an activity without being forced. In the context of games, autonomy can be promoted by providing players with more control over their actions. Another factor in motivation is *competence*, which fulfills the need to feel effective while also being appropriately challenged. Therefore, competence in SDT is closely related to flow theory. Lastly, *relatedness* refers to the need for social connections. Multiplayer games offer possibilities to interact with other players, fulfilling the need for relatedness [68].

Ryan et al. conducted studies examining the relationship between SDT and game motivation, successfully correlating the fulfillment of all three needs with intrinsic motivation in multiplayer games. In solo games, only competence and autonomy were associated with game motivation since there is no possibility for player interaction. Hence, exergames featuring a multiplayer option possess the greatest potential for sustained engagement. Additionally, exergames should emphasize higher perceived competence and a sense of autonomy to effectively motivate players [69].

2.2.3 Multiplayer in Exergames

The aforementioned novelty effect is a common problem for exergames and prevents successful long-term motivation for physical activities. Social physical play is a way to mitigate this problem. In a study comparing group and solo exergame play, participants who chose to play in groups played overall longer, had longer sessions and played more often than participants who preferred solo play [42]. Thus, group players had a significant higher adherence to play exergames. Social play relates to one of the pillars of the Self Determination Theory (SDT) [26] which refers to "relatedness" and fulfills the human need to belong by forming and maintaining relationships with others [13]. SDT explains intrinsic motivation by describing three psychological needs that humans want to fulfill. Therefore, incorporating multiplayer modes in exergames can lead to higher adherence for continuous gameplay. Furthermore, it has been shown that multiplayer exergames increase motivation [78] and engagement [50, 15].

Two types of social physical play can be implemented: competition and collaboration. In previous research positive effects on players have been observed for both types. Cooperation in exergames leads to higher self-efficacy, encourages pro-social behavior and kept players motivated to play long-term [49]. Competition, initially, increases the amount of effort that players put into playing exergames regardless of their tendency to be more or less competitive [61, 76, 49, 6, 75]. The effects of competition varies between players with more or less competitive personalities and can lead to negative effects in the latter while creating positive feelings for competitive players [76]. Competition can make less competitive players feel worse by decreasing their motivation and creating a negative mood [76]. It can lead to a decline of self-efficacy as well [48]. However, competition might still be a useful feature for players of all tendencies regarding competition as it encourages players to increase exercise intensity. It could be seen in Snyder et al.'s study that less competitive participants would increase workout intensity drastically when competing against an opponent compared to when they were exercising without any competitor [75].

While both competition and collaboration elicit positive effects in exergames, none of the two approaches is clearly favorable. Previous research has shown that multiple factors play a role when determining if exergames should implement competition or collaboration in a multiplayer game. Personality, game mechanics and other factors such as low or high level of skills can make one approach favorable over the other. In some studies, collaboration seemed to have a higher potential to engage players long-term in playing [49, 78]. In Peng and Crouse's study comparing competition, collaboration and solo play, competition seemed to result in high future play motivation and enjoyment [64]. However, participants were competing with a player in another room, while in collaborative play the players were sharing the room. Therefore, competition might have been preferred due to players feeling less restricted in moving around [64]. The effect of being less observed by other players might play a role in this scenario as well. A barrier for social physical activity can be anticipating negative reactions from peers [62]. However, in exergames attention is directed towards a screen which may help to feel less observed during physical activities and reduces body self-consciousness [79, 67, 12]. Moreover, it is not possible to see other players in real life while playing virtual reality (VR) exergames , which might increase this effect.

Interestingly, Shaw et al.'s results showed that less competitive players were not negatively affected by a virtual competitor. This was likely due to the fact that players were not competing with a real person and that the opponent was designed to match their skills which created a close game [73]. At the same time, less competitive players were increasing their workout intensity with a present virtual competitor [75]. Difficulty adaptation within a multiplayer exergame could further enhance self-efficacy as it provides support when needed and mitigate aforementioned negative effects in competitive games. Therefore, less skilled players have a chance to keep up with more skilled players. However, it is still to be researched if difficulty adaptation helps to overcome the problem of decreasing self-efficacy for players of certain skill levels or personalities. As it will be discussed in the next section, difficulty adaptation can lead to negative feelings it it becomes apparent to players in certain circumstances.

2.2.4 Balancing Skill Mismatches in Multiplayer Exergames

As exergames incorporate movement, players in multiplayer exergames might have mismatching physical abilities, which can worsen their experience. Some players might have less stamina, worse flexibility and coordination than their opponents. This might lead to frustration for less skilled players and boredom for more skilled players if games are mostly won or won by far by the more skilled player. Furthermore, previous research found that close games, when competing in computer games, are perceived as more fun than games where one player was far ahead [27, 87]. Providing assistance for players who are lacking skills can create a closer game and, therefore, a more fun experience for all players. This was already observed in studies concerning balancing exergames [39, 47]. More skilled players will face challenges that are more appropriate for their skill level, which is necessary for a better game experience according to Flow theory [25]. Less skilled players will have a higher chance in winning the game despite their abilities.

This opens up possibilities of accessible exergames, where able-bodied and disabled players can play together. Balancing skill discrepancies in exergames with players of large differing physical abilities have been investigated before. In a study by Gerling et al. able-bodied players were competing with players in wheelchairs in a dancing exergame. The dancing game was explicitly chosen for comparability as players have to make the same moves. Furthermore, it requires a combination of skills such as fitness and coordination. While the game assisted the player in the wheelchair, game balancing was not sufficient enough for a disabled player to win the game. However, disabled players enjoyed the assistance and did not mind losing to an able-bodied opponent, which they partially expected. Able-bodied players were mostly feeling guilty and did not perceive the game as fair [35]. Gerling et al. also investigated the self-esteem of players during gameplay in balanced exergames. Their findings show that weaker players experienced negative feelings and a decrease in self-esteem due to the explicit approach of the balancing methods. Furthermore, Gerling et al. assumed based on their findings that negative feelings arise when actual performance differ greatly with expected performance. It can be concluded that balancing of skills should be provided in a way that creates close games and adapt the balancing method for varying degrees of skills. Therefore, it can overcome various degrees of skill differences between players.

Mueller et al. provided a framework for balancing exertion experiences, which consists of four dimensions that an exergame designer should take into account [55]:

- Measurement
- Presentation
- Adjustment
- Control

Measurement: Balanced exergames will need to incorporate a measurement for exertion, which can be of effort or performance. An example for measuring effort is an exergame that converts heart-rate into movement in the game, which has been explored as balancing approach in previous research [77, 14]. Higher effort is, therefore, detectable as higher heart-rate. In both studies there have been cases of overbalancing less fit players as they reach higher heart-rates faster than fit players. Measuring performance would simply entail measuring the players performance within the game (e.g. how accurate targets have been hit).

Presentation: Exergames can be designed to present balancing either explicitly or hidden. Explicit balancing are clearly noticeable by players during gameplay. Hiding assistance can create the illusion that no players receive any assistance and play the exact same game. This can mitigate the problem of a decrease in self-esteem. Less skilled players can feel exposed when explicit balancing is used, which has been examined as effects of explicitly balancing exergames in previous research [35, 41]. It also affects how they perceive the value of a win as they could believe that they have only won due to the game's assistance.

Adjustment refers to the way how difficulty adaptation is implemented. Static balancing determines before game start who will receive assistance and how much of assistance they will receive. Instead of assisting the weaker player, handicapping of the stronger player is possible as well. In both cases, adaptation will not change during the game. However, roles of weak and strong player can change during game play as well as how much assistance or handicapping is needed. This can create overbalancing when a player becomes too strong due to the game's adaptation as seen in previous research [4]. Therefore, opponents will not be able to win. Dynamic balancing can solve overbalancing issues by detecting within the game when assistance is needed and, therefore, providing more tailored assistance.

Control: Furthermore, it has to be decided if players can control the balancing and how much they are in control of it. The opposite end would be a balancing approach fully controlled by the system. An example of player control is explicit selection of difficulty levels or the type of balancing by players. Offering players control can give them a sense of agency. However, the system needs to be designed to mitigate any possibilities of cheating.

Balancing differing skills have been already established in sports for a while. One approach is handicapping, which means that the person with better skills faces additional challenges. Altimira et al. investigated handicapping as a balancing approach in exergames and compared handicapping used in table tennis with handicapping in a virtual table tennis exergame [4]. The game either gave the weaker player a head start by adding points or the stronger player had to play with their non-dominant hand. Their results show that the skilled player experienced frustration as the added challenge made it hardly possible to win. Since they determined weak and strong players by a pre-study questionnaire as their balancing approach was static, they faced issues of overbalancing in some of the trials.

Since static balancing can create unfair treatment of more skilled players, Jensen and Grønbæk investigated dynamic skill balancing in an exergame that incorporates both virtual and real world elements [41]. Therefore, they could examine effects of physical balancing, implicit-digital and explicit-digital balancing. In their study they created a round based game, where participants took turns to hit targets with a ball at a large screen. They found that using a dynamic approach and giving assistance to both players depending on their performance could mitigate overbalancing. Their results show that while some participants preferred implicit (hidden) assistance as they felt less exposed as a weak player, explicit balancing was perceived positively by most weaker and stronger players in their study. Therefore, the preference for implicit and explicit assistance varies per player. Based on their results, they provide 4 design strategies for implementing balancing in exergames. While they have only used performance differences between players for adjusting the difficulty of the game, they firstly suggest to also include the individual performance of a player when determining the level of assistance. Secondly, they propose a combination of both explicit and implicit balancing or giving the user the choice of their preferred approach. As mentioned before, hidden balancing can help mitigating negative feelings of weaker players. However, they have also examined that players can receive an ego boost when their game has been adjusted to become more difficult. Thus, players became aware that they were doing well enough for the system to increase difficulty. As a third strategy, they state that players have to be aware of the assistance for higher acceptance and to not feel cheated on. Lastly, they propose that assistance should be available for all players, which mitigates overbalancing.

2.3 Virtual Reality

Virtual reality (VR) can be defined as the perception of being present in a virtual environment [80, 11] and the ability to interact with this environment [57] by using immersive technology such as a head mounted display (HMD) and controllers, both equipped with motion sensors. Wearing a HMD enables the person to change their view by moving the head. Thus, the virtual environment can be explored from any angle which leads to an immersive experience and the feeling of being present in this virtual environment. Apart from HMD's, other devices can be used to create a virtual reality such as Cave Automated Virtual Environment (CAVE) that projects onto the surrounding walls of the user. Therefore, the user does not need to wear any head mounted devices. VR technology has evolved from displaying VR on desktop displays to CAVE and HMD's [60], with the latter being now commercially available. Depending on intensity of the perceived presence and immersion, VR systems can be categorized into non-immersive, immersive and semi-immersive VR [10]. HMD's have been shown to create the most immersive experience [84] which can increase motivation for physical activity using VR [60]. Immersion makes VR an effective method to continuously encourage individuals to be physically active by playing VR exergames. Players can play in any kind of virtual environment which creates a unique experience. VR can create a more realistic experience, which games played on common displays are unable to do [43]. For example, the exergame Supernatural uses a variety of backgrounds depicting nature. Thus, the player feels like if they are training in different parts of the world instead of at home.

Nowadays, there are multiple VR exergames commercially available. Some of these are specifically targeted as fitness games that can be used for a home workout routine. Examples of these exergames are FitXR or Supernatural. Both of them require the player to hit targets with their hands and avoid objects by moving their body while either simulating boxing or dance moves. Many VR exergames exist which incorporate similar game mechanics but are advertised as entertainment games such as Beat Saver or Synth Riders. Aforementioned games already include the desired features such as multiplayer support, movement according to music and gradual acquisition of skills, which have been identified by Farič et al. as engaging VR game features [31]. Apart from these, there is a range of VR exergames simulating real world sports such as boxing and table tennis. These require the player to make similar movements as in real life. Previous research has also examined effects of VR paired with fitness devices such as exercise bikes, rowing machines and treadmills [85]. However, such games require players to have fitness devices at home. An example of a commercial game is Holofit, which offers the possibility to use fitness machines while playing.

Exergames are designed to be motivating and to increase physical activity. VR exergames, however, can offer additional motivating factors to enhance the players experience. The integrated motion sensors enable smooth tracking of motion and a natural feeling while using motion control. Realistic body movement during gameplay was also identified as one of the aspects for engaging VR exergames [31]. Exergames can be used to track players performance by tracking their movements to offer suggestions of areas to improve [57]. Due to its immersive character, VR can be used for simulating situations which Exergames can benefit from. For example, players can take roles in various positions during a sports game, which helps them to adapt to different positions and understand other perspectives [59].

2.4 Summary

Many definitions of serious games exist but authors agree on the aspect that serious games serve a purpose other than entertainment. For example, it can be used for education, awareness and motivation. One category within serious games are exergames that have the goal to encourage players to be physically active by playing games. For this, motion tracking sensors in controllers are needed to detect the players' movements. Flow theory was highlighted here as it explains how individuals can get engaged in an activity. Sports and games naturally have a high likelihood to induce a flow state as they can fulfill the necessary conditions for it. As exergames combine physical activities with games it has potential to motivate individuals to be more active. Multiplayer exergames can further increase motivation as well as physical effort since both collaborative and competitive exergames have shown positive effects on players. Flow theory has to be also taken into account when designing exergames to provide the ideal level of challenges for players. Furthermore, in competitive exergames the challenges of the game depend on the opponent's skills. If skills differ widely, it can lead to a game that is perceived too challenging. Thus, both players may be outside of the flow state. Competitive exergames can be balanced by difficulty adaptation in the form of providing assistance or handicapping. Previous research has mostly focused on static adaptation, where weak and strong player are determined beforehand and one of the players receives an adapted game, which does not change during gameplay. Results of these studies have showed that overbalancing occured in some of the trials. However, one study focused also on dynamic difficulty adaptation and could mitigate the overbalancing problem. To do so, it is necessary to adapt the difficulty during the game and to provide assistance to all players. However, none of the previous research has focused on exergames in Virtual Reality. VR offers the advantage that players are unable to see the real world. Hence, it is not possible to be observed by other players when moving. This can create a space where individuals feel less intimidated to be physically active.

3 Design Science Framework

This thesis will follow the methodology Design Science as described by Wieringa [90]. Design science research is concerned with designing an artifact and investigating this artifact in context and, therefore, combines research with design. As this thesis project will incorporate the design of a prototype and a user study in order to answer the research question, this methodology was chosen as research steps are integrated in the design process of the prototype. Before implementing a prototype it is also important to understand the requirements for this thesis' scope: rhythm exergames in VR. Therefore, a design process is needed as there are many ways of implementing dynamic difficulty adaptation. If designed badly, it can cause rejection of the game. Furthermore, the processes provide a logical order to conduct this research but leave it open to decide which specific research methods should be used within each process.

Wieringa proposes that processes of a design cycle and an empirical cycle should be merged. The design cycle will result in an implementation of an artifact, which will also be evaluated. The cycle iterates over its process steps in order to refine the artifact. It is concerned with improvement of the world by designing an artifact.

The empirical cycle defines the process steps to answer a research question. As opposed to the design cycle, the empirical cycle is concerned with answering knowledge questions and gathering knowledge about the world without improving it. Wieringa offers for both cycles checklists to follow during the research project. This simplifies the adaptation of this methodology.

3.1 Design Cycle

Part of the design cycle is to understand the social context of the artifact. This context includes the stakeholders who may use the artifact or are affected by it. In this thesis, stakeholders according to the design cycle are players of multiplayer exergames in VR. The design cycle is part of a larger cycle, in which the designed artifact is implemented and evaluated in the real world. This cycle is called engineering cycle and consists of the following processes:

1. Problem investigation
2. Treatment design
3. Treatment validation
4. Treatment implementation
5. Implementation evaluation

The design cycle is only concerned with the first three processes and iterates typically many times over these steps in order to refine the artifact. The artifact here is called treatment, as in Design Science the interaction between artifacts and context is designed to treat a problem. After the design cycle is completed, the treatment is implemented and evaluated in the real world. Figure 2 gives a guideline of questions that should be answered in each process.

The *problem investigation* consists of literature research and defining the problem. In this thesis, this was done as part of the literature research. Hence, similar solutions for balancing multiplayer exergames were reviewed as well as design requirements for designing engaging VR games, exergames and balanced exergames.

Within the *treatment design* process, the requirements for the treatment are defined and designs are created. Therefore, the requirements for the design of a dynamic difficulty adaptation will be defined. Part of this process is conducting preliminary focus groups and design focus groups. The first design focus group will result in multiple ideas for dynamic difficulty adaptation.

These will be validated within the *treatment validation* process by gathering the views of participants in another design focus group. Validation means here that it needs to be ensured that the treatment contributes to the goal of stakeholders. This goal here is to create a fair and motivating exergame between players of mismatching skills. This process always occurs before implementation as it predicts how the treatment will interact in the context but does not evaluate its interaction in the real world. A design theory is defined based on the results of this process, which describes a prediction of the interaction between treatment and stakeholders. Therefore, after conducting the last focus group, it will be analysed what effects each idea will likely have on players. Part of the design theory will also be the derived requirements from conducting focus groups. Based on this analysis, one idea will be implemented as a prototype.

The prototype will be implemented and evaluated in a user study. More iterations are not possible due to the time restrictions of this thesis. Evaluation is a different process to validation as it is concerned with investigating the interaction between treatment and stakeholders in the field. For this, it is necessary to have an implementation of dynamic difficulty adaptation in an exergame that players can play.

3.2 Methodology Design Cycle

In the following sections, the methodology of all four conducted focus groups will be described. Since preliminary focus groups and design focus groups differed in their procedures, they will be described separately.

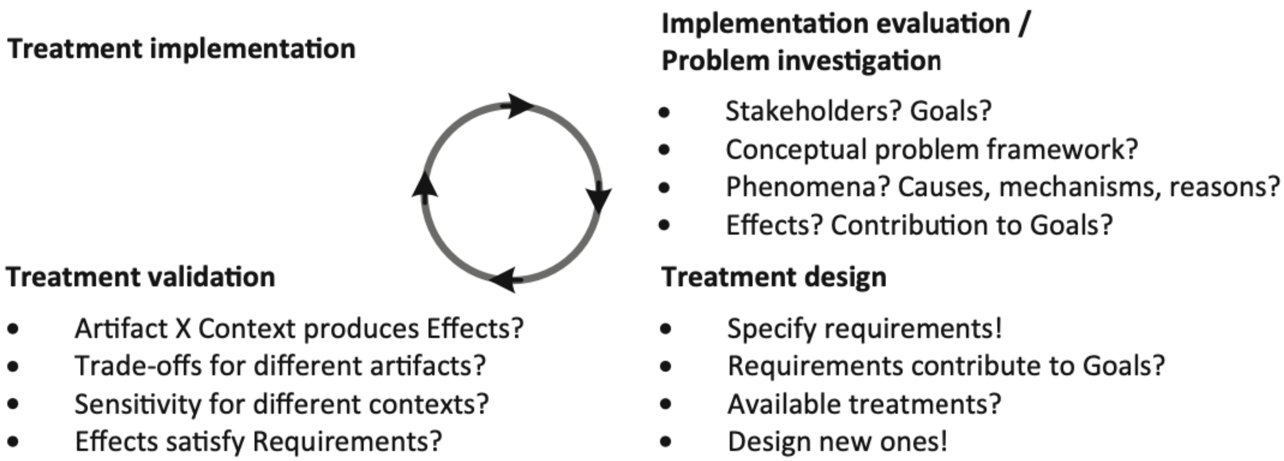


Figure 2: The engineering cycle [90]

3.2.1 Materials

For all focus groups a laptop and a smartphone were used to record audio. All participants were handed a consent form to sign and ball pens. Participants were shown a video of Beat Saber on the laptop. The first design focus group received an additional sheet with short descriptions of each requirement derived from the preliminary focus groups (see Appendix D.1). They were also offered additional blank paper sheets and pens when brainstorming and discussing ideas.

3.2.2 Participants

All focus groups except of one preliminary focus group were held with 4 participants. One preliminary focus group was held with 5 participants. All participants were recruited using convenience sampling. Participantss were either co-workers, student colleagues or friends of the researcher. All participants were either students or working full time at the time of the focus group. Recruiting was done either by directly asking co-workers and friends or sharing a sign up form in a Whatsapp group of student colleagues.

It was important that the design focus groups were held with participants, who have a background in designing User Experience UX), Human Computer Interaction (HCI) or game design. Thus, the first design focus group was conducted with 4 students. Three of them studied a Master's in Human Computer Interaction. The fourth participant studied a Master's in Game and Media Technology at the time of the focus group. The second focus group was held with two participants working as UX designer, one working as a software engineer and one participant with a background in Artifical Intelligence (AI).

3.2.3 Procedure Preliminary Focus Groups

Two preliminary focus groups were conducted to derive requirements for the development of dynamic difficulty adaptation in multiplayer exergames. Five participants, who stated that they were not competitive in general, took part in one focus group. The second focus group was then conducted with 4 participants who stated to be competitive in sports. This ensures that different ideas are gathered as it is likely that participants have different needs for difficulty adaptation depending on how competitive they are. Previous research has found that competitive exergames can induce negative feelings in less competitive players [76, 48]. Therefore, it is necessary to explore the different needs of competitive and non-competitive players.

The participants were able to try out a VR exergame prior to the focus group in order to get a feeling for the particular environment of VR games. The chosen VR game was Beat Saber as this thesis is focused on dynamic adaptation of dance rhythm games similar to Beat Saber.
Participants were encouraged to openly discuss questions in a semi-structured way. Thus, further questions can be asked during the focus group to ensure participants focus on the desired topics and to dive deeper into certain areas. The facilitator of the focus group had to make sure that every participant feels comfortable and had to encourage every participant to share their thoughts and opinions. The focus group started with a general question about how the participants would solve the presented problem of mismatching skills in multiplayer mode. During the course of the focus group more specific questions regarding dynamic difficulty adaptation were asked. Hence, dynamic difficulty adaptation was explained as it is the proposed solution in this research. The participants were asked to discuss the advantages and disadvantages of the proposed solution. More specific questions followed about the design of difficulty adaptation. Therefore, they had to discuss advantages and drawbacks of static vs. dynamic difficulty adaptation, increasing vs. decreasing difficulty and targeting only the strongest/weakest players vs. all players. The protocol of a focus group can be seen in Appendix 10.

3.2.4 Procedure Design Focus Groups

After conducting two preliminary focus groups, a third focus group followed with participants that have a background in Human Computer Interaction (HCI) or Game and Media Technology (GMT). The goal for this focus group was to use the requirements derived from the preliminary focus groups and create concrete ideas, how to implement dynamic difficulty adaptation in a rhythm exergame similar to Beat Saber. Therefore, participants were needed that have experience with designing games or technology according to users' needs. The participants were divided in pairs to create a low-fidelity paper prototype in two iterations. In total 4 participants, 3 with HCI students and 1 GMT student brainstormed ideas and discussed those in this focus group.

Participants were given a sheet listing the requirements. On this sheet, the most important requirements were highlighted. *Create close games* is important as it helps to keep players in flow by avoiding boredom and frustration. It is also important to *avoid demotivation of the less skilled player* when using explicit adaptation. Lastly, *make improvement visible* was highlighted as it was multiple times mentioned that improvement motivates and strong players are also rewarded with an explicit increase in difficulty. The requirement sheet can be seen in Appendix D.1.

In the first iteration participants were asked to design dynamic difficulty adaptation that is fair for all players. A paper prototype should be created by each pair. However, during the conduction of the focus group it became clear that the pairs needed time to understand the requirements and brainstorm various ideas. Therefore, they did not create a paper prototype. After the iteration the pairs were asked to present all ideas that they have brainstormed and discuss advantages and drawbacks. Participants where also asked which idea they preferred and for what reasons. In the second iteration, participants were asked to create an idea that takes into account the previous discussion and ideas of the other pair. The focus of this iteration was on encouraging players to move more during gameplay. This should mitigate the problem that players might play exergames with minimal physical effort. It was planned that the participants would discuss again ideas in pairs and design a paper prototype. However, this particular task required more technical knowledge about what movements are possible to track with a VR head-set. Furthermore, the participants struggled with creating initial ideas. Therefore, the participants were allowed to discuss in the whole group ideas that were presented to them by the facilitator. These ideas were: rewarding players with higher scores if targets were hit with a larger arm swing, rewarding players if they generally made more movements, providing players, who are lagging behind, the possibility to catch up by increasing the amount of movements. The participants also shared their own ideas, which were then discussed in the group. Lastly, the participants were asked how to personalize difficulty adaptation. The initial protocol for this focus group can be seen in Appendix 11.

After conducting the first design focus group, the ideas had to be evaluated. Therefore, a second design focus group was conducted, during which participants discussed the previous ideas and narrowed them down to one idea. Furthermore, it was discussed on how to combine this idea with ways of how to encourage movement during gameplay. At first, interviews were planned. However, the first design focus group included only four participants and thus, it was decided to conduct another focus group, which dives deeper into advantages and disadvantages of each idea. This approach allowed for a more comprehensive understanding of user requirements, facilitated discussions among participants, and aimed to determine consensus among participants regarding the identified game requirements.

3.3 Empirical Cycle

The empirical cycle is concerned with answering knowledge questions. These knowledge questions are the research questions that were defined in the previous chapter. While the goal of the design cycle is to create a solution for a problem and improve the world, the empirical cycle has the goal of advancing scientific knowledge. Wieringa provides a checklist with questions for the empirical cycle that can be followed while conducting the research [90]:

1. What is/are the knowledge goal(s)?
2. What is the current knowledge?
3. What will be contributed to the knowledge goal(s)?

As a first step, it is necessary to define the knowledge goal, which is the overall research question that should be answered. Afterwards, the empirical cycle has to understand the knowledge context of the artifact. For this, a literature research has to be conducted, which takes into account research findings of relevant areas. This directly relates to the problem investigation process in the design cycle. Finally, evaluating the implemented treatment will result in answers for the knowledge goal(s). This is part of the research problem analysis as can be seen in the empirical cycle (Figure 3.

Afterwards, it is necessary to design the research study by defining the procedure of the study, participants recruitment and measurements. Furthermore, the research design needs to be validated by providing justifications for the chosen methods and measurements and how the results can be used to answer the research question. The methodology chapter in this thesis includes both the research design as well as its justification. The research execution follows in the second phase of this Master's thesis, which then will lead to the data analysis and the answers of the research question.

3.4 Methodology Empirical Cycle

Main focus of the empirical cycle is to evaluate the prototype designed in the previous cycle. For this, a user study was conducted utilizing the implemented prototype to investigate the impact of dynamic difficulty adaptation. The hypothesis proposed that dynamic difficulty adaptation enhances players' perceived motivation, competence, and fairness, contributing to an overall elevated player experience compared to the identical game lacking dynamic difficulty adaptation. The following sections describe the necessary materials, participants' demographics, the procedure and the results of the user study.

3.4.1 Materials

For the execution of the user experiments, two Meta Quest 2 headsets, along with their controllers, were used. These devices were configured to accommodate developer

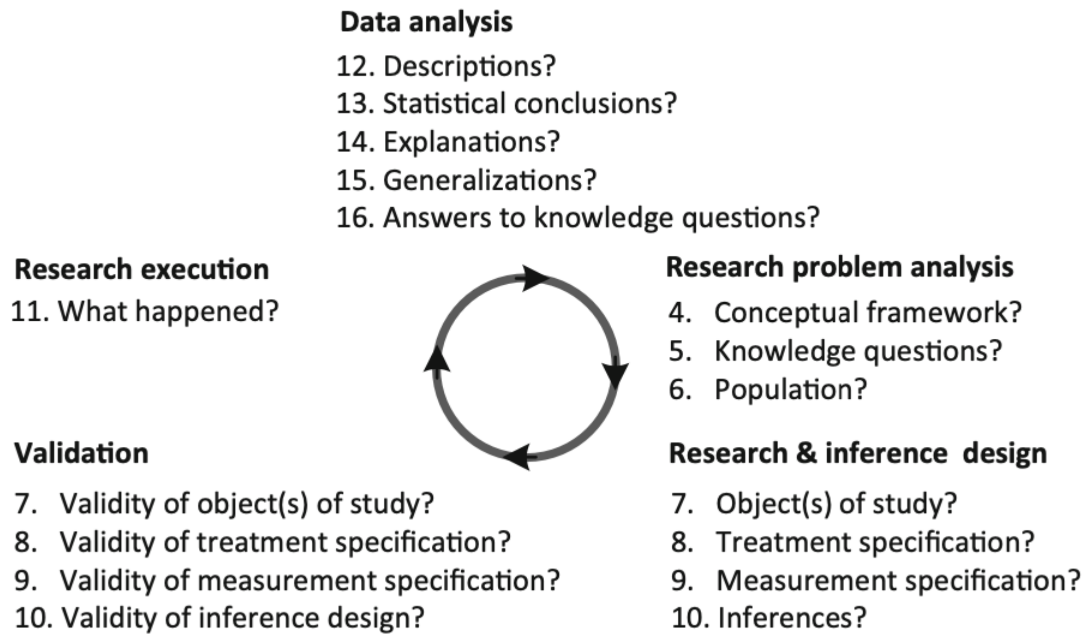


Figure 3: The empirical cycle [90]

settings. The prototype, developed using Unity, was deployed on both headsets via Unity. In addition, a survey was created using Qualtrics. Participants completed the survey on a laptop, with the option to use their personal devices or provided laptops. Heart-rates were measured using two smartwatches, while the recording of heart rates and other observations was documented either on paper or digitally via the Notes software on MacOS. Interviews were recorded using the app Voice Memos on an Iphone.

3.4.2 Participants

User experiments were carried out in pairs, which comprised a total of 14 pairs, thus involving 28 participants. Participants were recruited by convenience sampling, reaching out to friends, colleagues at work, and within the university for participation. Participants were required to play with someone they knew and, as such, were given the autonomy to select their gaming partner. This approach was chosen to explore potential differences in game experiences when playing against acquaintances, friends or family compared to strangers. It is expected that dynamic difficulty adaptation is most useful when playing against friends due to potential significant skill variations. On the contrary, playing against strangers was expected to result in fewer skill disparities, as skill matching mechanisms could be used in online game settings.

The participants' ages ranged from 23 to 42, with 9 (32%) being female and 19 (68%) male. The majority of participants had limited or no prior experience with dance rhythm games, as indicated by 11 (39%) having never played such games before and 8 (29%) having played once or twice. Additionally, 5 (18%) played such games rarely, defined as playing once every few months, while 2 (7%) played occasionally (every few weeks), and another 2 (7%) played often (weekly).

All participants had a background in Information Technology, either working for a software company or studying in the field. The majority of participants were employed in software development, with roles such as software engineer, data engineer, quality assurance engineer, software engineering manager, UX designer, and sales representative. Other participants included two PhD students, a recent graduate, and individuals working in marketing, office management, and product management.

Participants had to state their weekly physical activity and were categorized into high, mid and low activitiy. Twenty participants (71.5%) were highly active while 6 participants (21.5%) were moderately active and only two (7%) showed low physical activity.

3.4.3 Procedure

A between-subject study was conducted. This mitigates biases in the study results arising from groups that are fairly different in terms of physical abilities. Thus, such groups would be difficult to compare. Therefore, a participant pair played both conditions: (a) no difficulty adaptation and (b) difficulty adaptation is present. The effects of difficulty adaptation can then be compared. To mitigate learning effects and effects arising from exhaustion, the order of the conditions were alternating. At the start of the study, the thesis project was explained briefly to both participants. Thus, the participants were aware that the user experiment includes two games. One game included dynamic difficulty adaptation, while the other did not.

After the introduction, participants were then asked to start the survey on a laptop. The first page of the survey included further information about the user experiment and the thesis project. It also included a highlighted note about possibilities of motion sickness during the gameplay in VR. The participants were instructed to stop playing in case of motion sickness. The second page of the survey included a consent form with a checkbox to tick to ensure that the participant agrees. Once they have agreed, they received the smartwatches to wear during gameplay. Before the first game, the heart-rates of both participants were measured and noted.

The participants were explained basic interactions in VR such as how to press buttons. Afterwards, further explanations followed about how to play the game and its rules. Thus, it was explained that blocks should be sliced in the correct directions with the correct saber. The game also included a bonus system with two different types of blocks, which was explained as well. Participants were instructed to hit blocks with more force as this would increase their score. When participants stated that they understood the instructions, they put on the headsets. As both players were in the same room, the researcher could explain to both how to navigate to the game, which buttons to press on and which game they should play. There were two games to choose from: Game A (with adaptation) and Game B (no adaptation). Half of the 14 participant pairs started with game A and half started with game B to counterbalance biases. The games were identical except of game A would spawn more bonus blocks to the player that is lagging behind. Both games were about 4 minutes long. After starting the game, participants were asked if they could see each other within the game to ensure that they started the correct game and were aware of each other.

When the first game finished, the participants were asked to look at the smartwatch to check their heart-rate. For both participants, their heart-rate values were noted down. They were then asked who had won the game, which was noted down as well. The participants continued with the survey and filled out questionnaires determining their perceived competence, interest, effort and fairness.

Afterwards, the second game started, which was either the game with adaptation in case they played the game without adaptation as first game or vice versa. After the second game finished, the participants where asked to state their heart-rate again and who had won the game. These were noted down and the participants were asked to fill out the rest of the survey. This included questionnaires about perceived competence, interest, effort and fairness again. Furthermore, they were asked to fill out a questionnaire about their physical activity, social comparison and demographics.

When they finished filling out the survey, a short interview with four questions was conducted. Both participants were interviewed at the same time. The protocol of the user study can be seen in Appendix 13.

3.4.4 Independent Variable

The independent variable is the presence of dynamic difficulty adaptation. Thus, two players played a rhythm exergame at the same time in competitive mode with and without dynamic difficulty adaptation. After each game, dependent variables were measured via questionnaires.

3.4.5 Dependent Variables

To determine if dynamic difficulty adaptation for balancing exergames has positive effects, it is necessary to measure if players' self-efficacy increased while playing the game. Self-efficacy determines the players' perception of being able to win the game, which was measured as their perceived competence. Self-efficacy may increase for both strong and weak players as weak players may feel that winning the game can be successfully accomplished regardless of their opponent's level of skill. Additionally, difficulty adap-

Nr.	Statement
1	I felt capable while playing the game.
2	I felt I was good at playing this game.
3	I felt a sense of mastery playing this game.

Table 3: Competence scale: PXI

tation may boost strong players' confidence when they notice that the game becomes more difficult due to their good performance similarly observed by Jensen & Grønbæk [41].

Since difficulty balancing assists weaker players and increases difficulty for stronger players, it is important to see if both players still perceive the game as fair. In addition to fairness, it is also necessary to determine how motivated the players felt during the game. Players might feel more motivated due to the increased chance of winning or due to the fact that the games become closer, which has been positively associated in computer games in previous research [27, 87]. Furthermore, the players' end scores of each game will be measured to determine how close the game was.

3.4.5.1 Competence Measurement

For measuring competence the scale Player Experience Inventory (PXI) was used [1]. Particularly, the mastery and challenge constructs of PXI were chosen as these were the constructs that Abeele et al. conceptually mapped to the competence construct of Player Experience of Need Satisfaction scale (PENS) [1]. It was found that the mapped constructs highly correlated. PENS competence construct was also used in Jensen & Grønbæk's study that included dynamic difficulty adaptation [41]. However, PENS and other widely used scales such as Game Engagement Questionnaire (GEQ) lack empirical validation as opposed to PXI, which was validated over 9 studies [1]. The questionnaire was measured as a 7-point likert scale similarly to how it was measured in the validation studies.

3.4.5.2 Motivation Measurement

For measuring motivation two sets of questions of the Intrinsic Motivation Inventory (IMI) as 7-point Likert Scale [52] was used. The constructs that are of most interest in this study are *Interest/Enjoyment*, which directly measures motivation, and *Effort/Importance*. The latter is a useful scale to determine if players consciously tried to succeed in this game. Statements of both constructs are shown in Table 4.

Nr.	Scale	Statement
1	Interest/Enjoyment	I enjoyed doing this activity very much.
2	Interest/Enjoyment	This activity was fun to do.

3	Interest/Enjoyment	I thought this was a boring activity.
4	Interest/Enjoyment	This activity did not hold my attention at all.
5	Interest/Enjoyment	I would describe this activity as very interesting.
6	Interest/Enjoyment	I thought this activity was quite enjoyable.
7	Interest/Enjoyment	While I was doing this activity, I was thinking about how much I enjoyed it.
8	Effort/Importance	I put a lot of effort into this.
9	Effort/Importance	I didn't try very hard to do well at this activity.
10	Effort/Importance	I tried very hard on this activity.
11	Effort/Importance	It was important to me to do well at this task.
12	Effort/Importance	I didn't put much energy into this.

Table 4: Motivation scale: IMI

3.4.5.3 Fairness Measurement

As there is no standardised fairness questionnaire that has been used in research regarding gaming, a set of statements on a 7-point Likert Scale was created for this study. Similar research was done by Jensen & Grønbæk and Hwang et al. However, both have only included one question to determine the fairness of their implemented exergame balancing. Both have asked their participants if the game was fair as one question. The set includes both positive and negative (reversed) statements. Table 5 shows the set of fairness statements.

Nr.	**Statement**
1	I had a chance of winning until the end.
2	I was given an advantage over my opponent in the game.
3	My opponent was given a disadvantage over me in the game.
4	I was given a disadvantage over my opponent in the game.
5	I had a low chance of winning from the start.
6	My opponent was given an advantage over me in the game.
7	The game was close between me and my opponent.
8	The game was fair.
9	There was an equal chance of winning for both of us.

Table 5: Fairness scale

3.4.5.4 Interview

Both participants were asked questions in a short semi-structured interview. To understand how well participants perceived dynamic difficulty adaptation and how it can be improved, they were asked:

1. Which game would you prefer: the game with or the game without difficulty adaptation? Why?
2. What did you like about the difficulty adaptation? Why did you like it?
3. What did you not like about the difficulty adaptation? Why did you not like it?
4. Would you play this game again with difficulty adaptation? Why/Why not? If yes, in what situations?

Question 1 gives insights if dynamic difficulty adaptation was perceived well. The second and third questions were asked to understand what aspects should be included when designing dynamic difficulty adaptation and how it can be further improved. The last question explores in what situations participants would prefer dynamic difficulty adaptation.

3.4.6 Controlling Variables

Multiple controlling variables were measured as they may influence the dependent variables. The heart-rate was measured during the game to see if participants put effort into playing. Additionally, social comparison, demographic data and the amount of physical activity they engage in, were measured. With exception of heart-rate, all controlling variables were measured as part of the survey that participants filled out at the end of the user experiment.

3.4.6.1 Heart-rate Measurement

Participants wore a smartwatch to measure their heart-rate during the game. After each game, their heart-rate was noted down. Heart-rate was measured to see if participants put physical effort into playing. Besides the IMI questionnaire, high or low physical effort can be an indicator for how motivated players were during the game.

3.4.6.2 Weekly Physical Activity Measurement

Additionally, participants were asked to state how physically active they are during the week. For measuring this, the short version of the International Physical Activity Questionnaire (IPAQ) was used. While the long version was able to produce better estimates, the short version of IPAQ did not differ from its long version in terms of validity and reliability [24]. Thus, the short version was chosen as it is more compact. By measuring weekly physical activity, it is possible to see if there may be correlations between motivation, competence and how physical active an individual is.

3.4.6.3 Social Comparison Measurement

In order to understand differences in perceived fairness, motivation and competence between players, it was also measured how much a participant compares themselves to others. Players who compare themselves more than others may feel less competent when playing exergames with a stronger opponent. Thus, difficulty balancing may have better effects on participants with a high tendency of social comparison. To measure social comparison the Social Comparison Scale INCOM [36] was used. Furthermore, Schneider & Schupp proved validity and good model fits of their shortened INCOM version [71]. Thus, their proposed shortened version was used as can be seen in Table 6. Similarly to the fairness, motivation and competence scale the statements were measured using a 7-point Likert Scale. Table 3 shows the statements used for measuring social comparison.

Nr.	Statement
1	I always pay a lot of attention to how I do things compared with how others do things.
2	I often compare how I am doing socially (e.g., social skills, popularity) with other individuals.
3	I am not the type of person who compares often with others.
4	I often try to find out what others think who face similar problems as I face.
5	I always like to know what others in a similar situation would do.
6	If I want to learn more about something, I try to find out what others think about it.

Table 6: Social competence: shortened INCOM

3.4.6.4 Demographic Data Measurement

Lastly, the participants were asked to answer questions about demographic data such as age, gender and occupation and the type of relationship they have with the participant that they have chosen to play with (e.g. friends, family, colleagues). Furthermore, they were asked about their experience in playing VR exergames similar to Beat Saber. The demographics questionnaire was the last page of the survey that participants filled out during the user experiment.

3.5 Summary

This methodology adopts a design-empirical framework to address the research question, focusing on the process of developing dynamic difficulty adaptation for VR rhythm exergames while accommodating player preferences and personalities. The methodology describes two cycles. The first one focuses on participatory design and the second one focuses on empirical evaluation.

The initial phase consists of two steps preceding prototype implementation. In preliminary focus groups requirements for dynamic difficulty adaptation were derived. These requirements were then presented in a subsequent focus group, engaging participants in the creation of designs that align with their needs. This participatory design, inspired by Bodker's methodology [17], establishes a collaborative foundation for subsequent phases. The design focus group, that followed after, refined and enhanced the generated ideas. This resulted in the implementation of one idea as a prototype. The protoype will then be evaluated through a user study. The study, designed to collect quantitative and qualitative data provides insights into the practical implications and player experiences with the dynamic difficulty adaptation in a VR exergame.

In summary, the methodology integrates participatory design principles with an empirical evaluation cycle. This dual emphasis aims to not only ensure that the dynamic difficulty adaptation aligns with player preferences but also to provide empirical evidence regarding its impact and effectiveness in the realm of VR rhythm exergames.

4 Results Design Cycle

The first part of the thesis consists of gathering insights by conducting focus groups and interviews. These insights will be summarized and translated to requirements for the design of dynamic difficulty adaptation. Therefore, this chapter describes the focus groups' results including the derived requirements. Furthermore, one focus group was conducted to create multiple ideas, which were then further discussed with participants in a subsequent focus group. By taking into account the previously defined requirements and the insights from the interviews, one design was chosen for implementation.

4.1 Preliminary Focus Groups

4.1.1 Qualitative Analysis

For analysing the focus groups an emergent coding approach was chosen. The analysis followed the approach of the Straussian Grounded Theory, which includes open and axial coding [16, 22]. Preliminary codes were defined beforehand and directly relate to each discussion topic. Thus, previously defined codes were: preference for dynamic difficulty adaptation, static difficulty adaptation, bi-directional adaptation, handicap skilled players (uni-directional), decrease difficulty for less skilled players (uni-directional). Even though these codes were defined before, the aim of the analysis was to discover any relevant aspects that are important to consider when designing dynamic difficulty adaptation in VR exergames. The recordings of the focus groups were transcribed and either assigned to one of the preliminary codes or a new code was created (open coding). The criteria codes were then grouped, which formed categories (axial coding). The most important codes were determined by the number of mentions by both groups. The categories and their most important codes will be further explained in the following sections. Requirements were then formulated based on the most important codes. In Appendix C.1 all found codes and derived categories as well as the number of mentions can be seen.

4.1.2 Results of Both Groups

This section explains the most important categories and codes found in both groups. Overall, both groups mentioned that seeing improvement is important to continue playing exergames. Players should also have control over the type of adaptation. For example, there could be different modes offering dynamic or static difficulty adaptation. Both groups also agree that close games, where all players are more or less performing similarly in the game, are the most fun. This aligns with previous research that found that close games are more enjoyable when playing computer games [27, 87].

4.1.2.1 Motivators to Play Exergames

It can be motivating to play against higher skilled players, especially if players have a competitive personality. It was mentioned that winning can give positive feelings but is not always necessary. Oftentimes seeing the improvement over time is more motivating than winning games. However, failing every game as well as being very far behind can be demotivating. Thus, players should think that they have a chance to win. Both groups mentioned that the game could incorporate rewards such as game items for returning to the exergame after a while.

> P3: If you play against individuals that are on a significantly higher level than you, it actually helps you to excel quite a lot as well. But I think that depends on how competitive as a person you are. If that drives you or demotivates you.

4.1.2.2 Disadvantages of Dynamic Difficulty Adaptation

Both groups mentioned more disadvantages of dynamic difficulty adaptation than advantages. It was feared that the system will overbalance and, therefore, making it too easy for the weaker player and too difficult for the stronger player. Additionally, dynamic difficulty adaptation could facilitate cheating. For example, players could play badly on purpose in the beginning to receive an advantage.

> P3: I feel like if you have VR boxing, for example, and you have one player putting the whole body in and just making giant moves and just burning lots of calories and the other person, for whatever reason, maybe can't even stand up and just sitting and making small movements. It would feel unfair, if they would get the same score.

The quote also makes clear that increasing physical effort in an exergame should reward the player. Fairness was another mentioned point as players could feel unfairly treated if they put a lot of effort into the game and do not get rewarded more. In addition, a game should be predictable. This means that if a game is played repeatedly, it should stay the same without changing the difficulty given that the game offers different levels of difficulties. Another concern was mentioned regarding less skilled players who may feel demotivated if they are aware of a decrease in difficulty as it means that they are performing badly. Participants in previous research have experienced this when using explicit balancing of exergames [35, 41].

> P1: Maybe if I'm super conscious that it has changed and it's kind of obvious then I don't know how good that would make me feel. [...] I might feel a bit demotivated. [...] If the change was subtle, it might work in a motivating sense.

4.1.2.3 Adaptation Needs to Be Tailored to Players and the Game

Participants agreed that the usefulness of dynamic difficulty adaptation depends on the players' personalities, the particular game and implementation of difficulty adaptation and the context of the game. Therefore, competitive players may prefer the challenge of an opponent who is more skilled than them. Mario Kart's game mechanic that assists players on the last place while punishing the player on the first place was seen as a positive feature that creates a more fun experience for everyone. Players might also prefer to use dynamic difficulty adaptation if they prioritise fun over competition. Participants suggested that games should offer dynamic difficulty adaptation as an additional mode to select. Alternatively, players might want to change settings to create a difficulty adaptation tailored to their needs with the possibility to decide how much they want to be assisted.

> P5: So some individuals might just do it casually and then they just want something that they enjoy and some individuals might want to improve continuously and then it will up the difficulty as they get better.

4.1.2.4 Advantages of Dynamic Difficulty Adaptation

Certain advantages of dynamic difficulty adaptation were mentioned that can lower the probability of the occurrence of the aforementioned disadvantages. Most importantly, both groups would enjoy close games. According to participant P3, "everybody gets more dragged into the middle", which creates a game where players can still catch up and the outcome of the game is only determined at the end. As a positive example for a commercial game, that incorporates dynamic difficulty adaptation, both groups mentioned Mario Kart. The game includes a game mechanism that makes it difficult to predict who will win as the placement within the race can change drastically until the end. Participant P4 states that by adjusting the difficulty of all players, the game can become more fun, given that players primarily play for fun rather than competing in a serious context such as in athletic competitions.

> P4: So you play against each other, but it's not about who's going to win the title. [...] And then you can play this mixed mode where it's keeping both of them close. So for the weak person it gets easier. So it gets to the middle. For the more experienced person, it gets more difficult, so it gets to the middle and then most individuals have fun. But it's just about playing.

Furthermore, games could be designed without explicit levels while internally the game adjusts the difficulty accordingly to keep every player in the flow. Therefore, players may not have the need for consistency in games as they are not biased by explicitly shown difficulty levels.

> P1: I feel like it's not so conducive to have really explicit levels and then I kind of like that the game then will just dynamically adapt to whoever is

playing in the moment to the difficulty that they are suited for without you already kind of pre-selecting. [...] You just put in a game and then you just kind of slowly have to figure out, what's my level increasing. And then [...] you're not so biased before you go into the game.

Dynamic difficulty adaptation also allows for gradual difficulty increase and decrease. There might be cases that the provided discrete difficulty levels are not sufficient as a player might need a level in between two levels. Gradual adaptation can provide an ideal level of challenges for each player.

> P6: I think that the benefit of the dynamic is that the change is happening gradually while with static, it's like easy and then medium or difficult. It's like the steps might be a bit too high.

4.1.3 Differences between Competitive and Non-Competitive Groups

Competitive group: The competitive group was more open to accept dynamic difficulty adaptation in games than the non-competitive group. They discussed more examples of how to implement dynamic difficulty adaptation. Their biggest issue with this approach is the potential for creating an unfair game for skilled players as skilled players expect that their higher skill will lead to higher rewards. However, they also mentioned that this would be likely only a case for very competitive players. Thus, if players only play for fun with friends, it would not necessarily be a problem that less skilled players receive assistance or skilled players are handicapped. Dynamic difficulty adaptation should not be used in actual game competitions nor for experienced players, who are trying to improve their position in a global leader board or similar. However, since the goal is to create multiplayer exergames that motivate players to be physically active regularly, it targets players whose focus is not to participate in serious competitions, but to adapt a healthier lifestyle.

> P7: You can have like casual mode, let's say, where you just fight each other and best player wins, but it's all good fun. Then you have competitive mode with rankings and what have you. That might, you know, be weird if you're playing against someone and suddenly, you know, that person has boosts or aim assistance [...] but that's unfair.

Non-competitive group: The non-competitive participants discussed more potential disadvantages of dynamic difficulty adaptation. They preferred static difficulty adaptation over a dynamic approach similar to how it is present in current games: Players choose an appropriate level and get matched to other online players with the same level. However, this is not possible in the scenario where players compete against friends with higher skills.

The non-competitive group discussed the VR space as well, which was not mentioned by the competitive group. It offers a unique experience which enables players to view different virtual worlds with shared information between players. Thus, players might not be able to see if other players receive an increase or decrease in difficulty. It would also be possible to create a virtual world where players believe to play the same game even if their game was adapted differently and therefore differs.

> P1: So you're constructing kind of like a different reality, like you're not actually looking at the person that is necessarily better. Are you actually going to compare yourself to the person in the virtual space as opposed to in the real space? So why not change the rules? [...] Maybe it's not as obvious that one person is better because you're not actually engaging with them in real life. [...] So you can kind of build or do whatever you want in some sense and just fool individuals into thinking that there's true competition.

4.1.4 Requirements for Dynamic Difficulty Adaptation

The results of both focus groups were taken into account to develop the following requirements for dynamic difficulty adaptation in VR multiplayer exergames. The found codes in the categories *requirements*, *motivators* and *advantages/disadvantages* have been translated to requirements. These requirements summarise aspects that need to be taken into account for designing a dynamic difficulty adaptation in a rhythm dance game such as Beat Saber. Only requirements for game play are listed as the design focus group will create ideas for adapting the game play. However, other requirements, that will not be further explained here, were mentioned as well such as: offer modes and settings for adaptation, make clear that players play an adapted game and reward players to play exergames again over a long period of time.

4.1.4.1 Make Improvement Visible

As improvement is more important than winning games long-term, the game can incorporate feedback on how players improved. Furthermore, stronger players get motivated by seeing an increase of the difficulty due to their performance. Thus, difficulty can be increased explicitly to make players aware that they are performing well. Thus, explicit difficulty adaptation is one form of making good performance visible and therefore also improvement. Positive feedback regarding improvement satisfies the need to feel competent, a crucial factor contributing to intrinsic motivation according to self-determination theory.

4.1.4.2 Create Close Games

The game is ideally close to the end. Thus, players, who fell behind, have the chance to catch up and the player on the first position can always be challenged by other players. This also helps to keep players engaged according to Flow theory and avoid boredom or frustration. This means that no player receives an adaptation that makes the game too easy or challenging.

4.1.4.3 Do not Demotivate Less Skilled Players

Less skilled players should not get demotivated when their difficulty visibly decreases due to bad performance. This can decrease their perceived competence, which in turn leads to lower intrinsic motivation. This can be avoided by creating subtle adaptations. Another way is to let players decide if they want to use a temporary advantage in the game. This means that the game rather offers the possibility of assistance than changing the difficulty immediately without the player's input. Therefore, the player stays in control of the adaptation, which also contributes to a feeling of autonomy.

4.1.4.4 Add the Right Amount of Randomness

Chance can be introduced to make the game more unpredictable. Thus, the game outcome is not determined until the end. It can help players feel as if there was still a possibility for the game's outcome to turn around. However, too much randomness can be frustrating for players when putting effort into the game.

4.1.4.5 Create a Unique Virtual Reality for Each Player

In virtual reality it is not possible to see the real world, which can be an advantage in exergaming. Players do not necessarily see the physical activity of other players. Thus, a virtual reality can be created where players have a common ground, but their virtual realities differ. Players do not necessarily see if someone has an increased/decreased difficulty unless it is the intention of the game to share this information.

4.2 First Design Focus Group

4.2.1 Qualitative Analysis

Similarly to the preliminary focus groups, the audio recording was transcribed. Transcriptions were then analysed using an emergent coding approach based on Straussian Grounded Theory [16, 22]. Codes were created based on statements (open coding), which were then again grouped by similarity. The resulting codes were then categorised in four categories (axial coding) that directly respond to the four discussion topics of the focus group. These categories will be explained in the following sub-sections.

4.2.2 Results

The analysis of the second design focus group discussions revealed four main categories. These sections will cover specific aspects highlighted by participants, which were most frequently discussed.

The first part focuses on what participants considered necessary for adjusting the game's difficulty during play. It explores their ideas on maintaining fairness in the game, especially when the difficulty changes dynamically.

The next section highlights the main points on how to encourage movement in a rhythm dance game similar to Beat Saber. The final section assesses which game idea participants preferred and the reasons behind their choice.

4.2.2.1 Adaptation Requirements

Participants discussed requirements that are needed to achieve a good acceptance of dynamic difficulty adaptation. Three requirements were most important. Firstly, the group agreed that they would like *transparency* when games of players are adapted and therefore result in varying games for each player. Thus, creating games that seem to be the same but differ for each player were rejected. As a reason, it was stated that players might feel cheated by the game. For example, skilled players might perceive it unfair to lose if they cannot see what has changed in the opponent's game. Therefore, transparency may lead to higher perceived fairness.

Another requirement is *player control*, which means that any game advantages that a player receives should be activated by the player. Thus, the focus group rejected the idea that difficulty decreases automatically for players. This adds another type of interaction within the game and leaves the decision to decrease difficulty in the hands of the player. Therefore, dynamic difficulty adaptation is more accepted, as players have actively decided to use the advantage that they were offered. A possible example for an implementation in a game similar to Beat Saber could be that the game spawns special cubes that activate bonuses when the player manages to hit them.

The third requirement is *subtle changes*. The participants would prefer subtle increases or decreases in difficulty to avoid that a player gets overbalanced. Hence, no player should get an advantage or disadvantage that is too big. Ultimately, the goal is to create close games which ideally give possibilities to catch up and win until the end for both players.

4.2.2.2 Fairness Requirements

In the first task, participant pairs were asked to create a dynamic difficulty adaptation that is fair for all players. Thus, in the discussion requirements were mentioned that may create a greater perception of fairness. Participants would like to see their opponent as an avatar. The avatar does not have to be realistic but it should be visible how the opponent is moving in the game. This can create a fairer game according to the participants. When players see that their opponents moves a lot and put effort into playing even if they are less skilled, a loss might be more accepted.

> P1: And maybe it also ties into like the fairness part a bit that if you see that, okay, like your opponent had like bigger blocks or like clearly an easier

> game, but you still see that they put in the effort and they were moving a lot and they were trying then maybe it still feels fair kind of.

Furthermore, they would like to see their friend, who they are playing with, as an avatar simply to feel more connected to them. Otherwise, it could create the feeling that players are playing each individually and only scores are compared at the end.

Being able to see achievements can lead to a fairer game when these achievements highlight how a player has improved over time and multiple games. One participant mentioned that dynamic difficulty adaptation can lead players to believe they are better or worse in the game than they really are. Therefore, showing their achievements can make them aware of their true skill level.

Another requirement is that all players' games will be adapted. Therefore, a skilled player can receive advantages in the game as well given that they are falling behind. Therefore, the adaptation will be perceived as fair since all players could benefit from dynamic difficulty adaptation.

4.2.2.3 Encourage Movement

Multiple ideas to encourage movement in VR exergames were discussed. The participants stated that they liked the idea to encourage movement by incentivizing effort. For example, the game could increase score if players move more during the game.

> P1: And that ties a bit into [...] the whole adaptation and [...] that it could also be that [...] weaker players get rewarded not just for like, let's say, how many targets they hit, but also on the effort that they put in. That could be measured by maybe like the amount of movement or the intensity of the movement, the speed.

As stated by a participant, incentivizing physical activity in general, could provide an additional way for less skilled players to score higher. The participants discussed that complicated patterns and poses, which players have to follow additionally to hitting blocks, adds humour to the game and, therefore, increases entertainment. This was the most discussed idea and can be combined with the idea of placing game elements further apart. Furthermore, they discussed jump scares, consisting of surprising and shocking elements to create a sudden increase in their heart-rate. However, it was the least liked idea.

4.2.2.4 Best Ideas

Three ideas where discussed most and stated as favorite ideas. Firstly, power-ups as a way to add bonuses and introduce ways of simplifying the game for a specific amount of time were seen positively. Another idea, which participants liked, was the possibility to attack the opponent. Therefore, instead of decreasing the difficulty of the player, the opponent's game difficulty increases. Participants were in favor of this idea because it added interactions between players as stated by a participant:

> P2: The actions you're able to get, especially to mess with other people, I think, would make it more interactive and make it more fun as well for me.

Rhythm dance games often do not include ways to directly interact with opponents as players have to follow the same dance moves, steps or movements and receive a score depending on accuracy. Lastly, movement patterns and poses were among the participants' favorite ideas as they add humour and the possibility to encourage players to move more during gameplay.

4.3 Second Design Focus Group

4.3.1 Qualitative Analysis

As previous focus groups the qualitative analysis followed the methodology of emergent coding based on Straussian Grounded Theory [16, 22]. For this, audio recordings were transcribed then coded. Through open coding, new codes were created. The resulting codes were grouped to five categories (axial coding): favorite idea, requirements, ideas to activate power-ups, reasons for choice, additional features.

4.3.2 Results

Following results summarise the most liked ideas of the participants, reasons for their choice and further ideas for enhancing a rhythm dance game in VR. The participants were also asked to discuss concrete ideas how their favorite ideas could be explicitly implemented.

4.3.2.1 Favourite Idea

The participants favored a specific idea that emerged during discussions — a combination of power-ups and attacks designed to influence opponents' game play. The idea of being able to disrupt an opponent's game resonated strongly as it encourages interactions between players. This idea was perceived as both fun and original.

> P1: And that's again a lot more interactivity right. That's because then you're trying to prevent the other [player to score].

Power-ups alone were seen to be less interactive, as the opponent would not notice if the player received an advantage.

> P2: But with those power ups, it's only for you. So, the other player cannot see that you are having that advantage.

Moreover, participants proposed a mechanism where power-ups are activated through increased movement during game play, visually represented by a loading bar. This loading bar advances more rapidly for players falling behind. Thus, less skilled players unlock bonuses such as power-ups and attacks faster than their opponent and therefore, receive more opportunities to catch up. The interactive dimension of attacking opponents and increasing their difficulty adds unpredictability to the gaming experience.

Combining power-ups with attacks seemed fair to participants because it provided a viable alternative for weaker players to succeed, requiring less skill but incentivizing physical activity. Furthermore, the idea was seen as a means to challenge stronger players, adding an extra feature for competition.

> P3: I think the issue is [...] being worse than another person, missing the beats or something because it's all like rhythm based. Right? So maybe you can get like if you're messing with someone's game, it's like more fun and less a little bit about the rhythm part.

As development progresses, this preferred concept, which combines power-ups and attacks, will be transformed into a playable prototype of a VR game. Aligned with participants' preferences, the implementation phase will refine the integration of power-ups and attacks, aiming to deliver an engaging gameplay experience that aligns with the requirements that were derived from this and preceding focus groups.

4.3.2.2 Requirements

The participants mentioned several requirements to implement the favored idea. Firstly, to ensure that the game is perceived as fair, players should not be given advantages without putting effort into playing. Thus, the game should offer opportunities to catch up, but in order to use these opportunities, players will have to actively do something. For example, players need to hit targets or move more. One participant mentioned that they would rather incorporate strategic elements to the game, which enables players to catch up, than simply pressing a button to get ahead:

> P3: I really like to keep it on the technical part of the game where you [...] use these technical aspects to gain an advantage instead of [...] pressing a button and getting like a power up [...].

Furthermore, another participant mentioned that the power-ups and attacks should not be overpowering. Mario Kart was mentioned as an example, where power-ups can overpower players.

> P2: It's hard to find the balance of like being overpowered [...] like a bullet. [...] I was riding first the whole game and then you get like one thing and you're [thrown to the back]. But it's like, make it so your own efforts rewards you for your, uh, catching up.

When attacking the opponent, it should still be manageable for the opponent. For example, the opponent can block attacks or end them earlier by performing certain movements.

> P3: I think being able to like remove [the attack] by doing something so not that the other person does something and you are stuck with it for like 30 seconds. You having the chance to like get rid of the [attack].

Furthermore, one participant mentioned that the game should be self-explanatory to quickly learn game mechanics.

4.3.2.3 Further Ideas

More ideas to further enhance the game were mentioned, which were not directly connected to the implementation of difficulty adaptation. Among these ideas was getting bonuses for streaks of hitting blocks correctly. This is similar to Beat Saber's game mechanic, where the score is increased when streaks are activated. Participants also mentioned that a calorie counter would be beneficial. The Meta Quest 2 headset, which was used for the user study, has incorporated a calorie counter. It can be seen by looking to the ceiling and vanishes when looking down again. Furthermore, it was mentioned that they wish for strategic elements in the game, hints that point out to players how to catch up, and feedback about who is currently leading.

4.4 Requirements

Table 7 summarizes the requirements derived from all 4 focus groups. Note that similar requirements have been merged together. *Subtle adaptation* and *Do not mess too much with opponent* were merged together as the latter is a direct consequence of subtle adaptations. Furthermore, *Player in control*, *Player needs to put effort to get power-up* and *Avoid demotivation* were merged as demotivation can be avoided by putting players in control. *Create a close game* and *Provide a way to catch up* are directly related and therefore, merged. A short overview of all requirements derived from the focus group results can be seen in the Appendices C.1, D.2 and D.4.

Nr.	Requirement	Description
R1	Create close games	In order to create a fair game without frustrations, it is important to balance scores between players during game play. Ideally, scores differences do not increase greatly or will be quickly balanced out after increasing.
R2	Avoid demotivation of the less skilled player by player control	By offering opportunities that players have to actively take, they are put into control. This means that they can choose to not use the help that they receive and therefore, they may feel less exposed as weak player.

R3	Make improvement visible	Improvement can be made visible after the game ends. For example, information about their true performance can be given later.
R4	Add the right amount of randomness	Randomness can make games unpredictable and more interesting. However, it should not include too much randomness, as players may feel the outcome is unpredictable and detached from their effort and performance.
R5	Make adaptations transparent	Players should be aware that they play an adapted game and should be able to notice when they receive assistance due to their weaker performance, as they do not want to think that their performance is not solely based on their true skills.
R6	Subtle adaptations	It was mentioned that attacks on opponents should not be too extreme, as they should not feel disadvantaged over the player who receives assistance.
R7	Seeing opponents effort through avatar movements	Participants would like to see their opponents in a multiplayer game feel like they are playing with another person rather than alone for themselves. They also want to see their opponent if they are interacting by attacking each other.
R8	Both players need to get positive adaptations	To create a fair game, both players should be able to receive assistance when they are lagging behind even if they were previously leading.

Table 7: Requirements for dynamic difficulty adaptation

4.5 Summary

Four focus groups were conducted to explore dynamic difficulty adaptation in the context of VR dance rhythm exergames. The initial two focus groups involved general discussions on dynamic difficulty adaptation, including its advantages and drawbacks. Subsequently, the following two focus groups centered around the specific design of dynamic difficulty adaptation within a rhythm dance game with similar game mechanics as Beat Saber.

The first two focus groups generated a list of requirements crucial for designing dynamic difficulty adaptation in exergames. These requirements served as the foundation for subsequent design-focused focus groups. The first design focus group aimed to generate ideas aligning with the identified requirements and creating a close game that encourages physical activity. These ideas were then subjected to evaluation and refinement in the second design focus group.

The concept of incorporating power-ups and attacks emerged as the favored idea. This concept not only introduces interactions between players, but also adjusts the difficulty for both skilled and less skilled players. While power-ups decrease the difficulty of the game for the player lacking behind, attacks increase the difficulty of the opponent's game. The activation of power-ups and attacks is tied to physical activity. Importantly, any player, who falls behind, is subject to the adaptation. This approach is perceived as fair, ensuring that players who fall behind receive more bonuses regardless of their previous performance.

The chosen idea, a combination of power-ups and attacks, will be implemented in a playable VR prototype.

5 Prototype

For conducting a user study, it was necessary to implement a playable prototype. Since this prototype is intended for use with a VR headset, it was developed in Unity version 2021.3.16f1, utilizing its XR (extended reality) framework along with the XR Toolkit to facilitate the integration of VR interactions. For writing Unity scripts, Visual Studio Community 2022 version 17.6.5 was used. The game mechanic is similar to the popular rhythm dance game Beat Saber, known for its game mechanics where players strike blocks with differently colored sabers.

The implemented game was designed to be a simplified version of Beat Saber. Additionally, it introduces unique game elements absent in Beat Saber. These elements were tailored in accordance with the requirements and findings of previous focus groups. Approximately two months were dedicated to the creation of this prototype.

The prototype features a multiplayer mode realized through the Photon Pun library. This library establishes a network that seamlessly connects multiple players within the game. The game was designed to be playable by two players. Players can see each other within the game (see Figure 6).

5.1 Game Mechanic

Similar to Beat Saber, blocks are sliced through by sabers of different colors in various directions. The bonuses spawn when a loading bar is fully filled. The loading bar can be seen in the background of Figures 5, 4, 7 and 6. The more a player moves, the faster the bar fills up. Once it is fully loaded, it will spawn two bonus blocks. Thus, the more someone moves, the more bonuses can be unlocked. The bonus blocks are spawned simultaneously on the left and right sides. When one bonus block is hit, it activates a power-up or attack for 10 seconds. Therefore, bonuses are only activated if a player sliced them. Bonus blocks are also spawned further away from normal blocks, making it harder to hit them simultaneously with normal blocks. Therefore, players still need to put effort to activate bonuses, which is aligned with the requirement *R2: Avoid demotivation of the less skilled player by player control.*

The game does not include a lot of randomness as aligned with *R4: Add the right amount of randomness.* However, the vertical position of the bonus blocks is spawned randomly. Furthermore, bonus blocks can be spawned simultaneously when many or few normal blocks are spawned. Therefore, hitting bonus blocks can be harder or easier depending on the timing of them appearing. Players cannot choose when bonus blocks are spawned as they are activated when a certain movement threshold is reached.

The power-up block with a star icon decreases the player's game difficulty, for instance, by multiplying points by a factor when a block is successfully sliced. It can also make the game easier by allowing the player to hit blocks from any direction. The non-adapted game will only include the mechanism of multiplying scores with a factor. The adapted game will have different mechanisms depending on how large the difference in score between players is. Furthermore, movement is incentivized by multiplying scores by a higher factor when blocks are sliced with greater acceleration. The factor is directly related to the amount of acceleration. For example, if a block is hit correctly with the least amount of acceleration possible, the player receives 10 points. However, with more acceleration, the player can receive up to 100 points. If players miss blocks or slice them incorrectly, they will receive minus points, which pop up in red after slicing a block. Therefore, it is easier for players to see if they hit a block incorrectly.

The attack block with a sword icon, on the other hand, attacks the other player. Thus, the opponent's block size decreases, blocks rotate for greater difficulty, or their score is divided by a factor. The factor is determined by how often the other player hits an attack block within 10 seconds. The factors by which their score is multiplied or divided are visible to players, indicated on the block (see Figures 7, 5). Similarly to power-up blocks, for the non-adapted version an attack will only decrease the opponent's score by a factor. The adapted version will have different mechanisms depending on the score difference.

Additionally, similar to Beat Saber, a player hitting 10 blocks correctly in a row receives a multiplier added to their score, doubling with each successive set of 10 blocks hit in a row. This multiplier is written as text above the loading bar for each player. When a player misses a block, the multiplier will become the default value of 1, meaning the score will not increase.

The players are standing next to each other within the VR game and are visualized as avatar with a head, torso and arms holding sabers. Thus, players can see each other moving if they look to the side. This aligns with *R7: Seeing opponents effort through avatar movements* as they can look at each other during the game. The blocks are spawned in front of each player. Therefore, players can see their opponent's blocks next to theirs. A gap between the players' positions indicates which blocks correspond to which player. The player who first enters the game is always spawned on the left side of the VR room. The second player spawns on the right side.

Certain game mechanics were mentioned as features in focus groups, including the addition of streaks when a player accurately hits multiple blocks and the incentivization of movement, such as achieving a higher score by hitting blocks with more acceleration. Furthermore, it was decided to not implement requirement *R3: Make improvement visible* as this requires an accurate comparison of a player's past performance with their current performance. The prototype will only be played for a short session during the user experiments. Thus, past performance does not exist at that time.

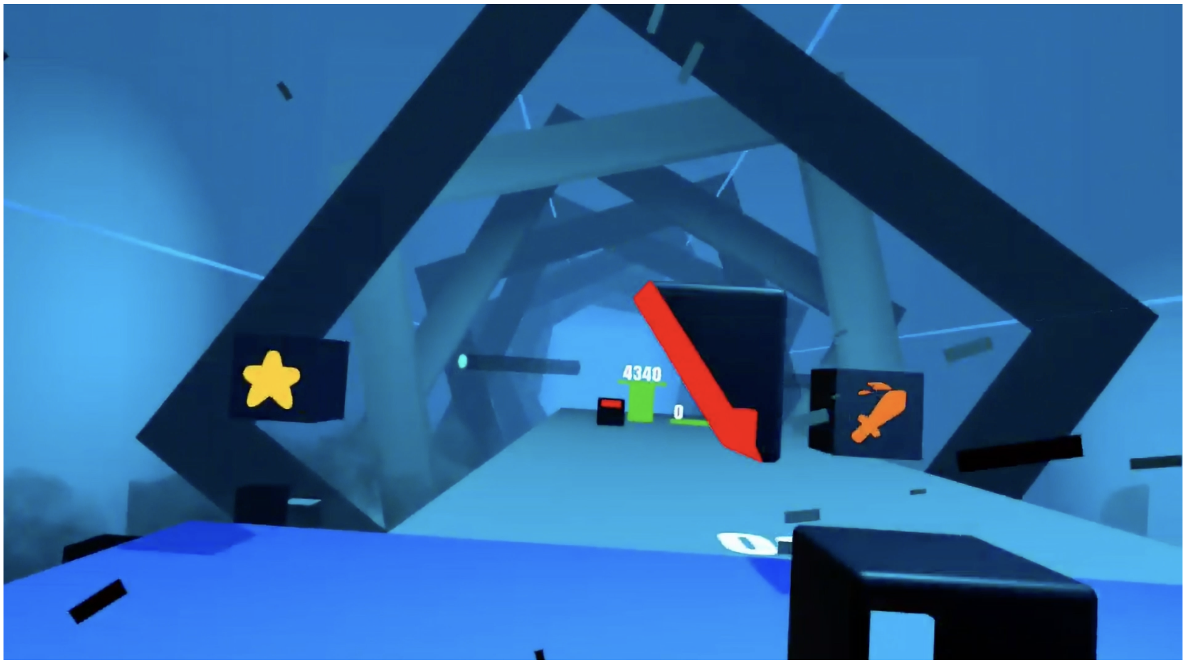


Figure 4: Exergame showing bonus blocks

The game included visual implementations of factors and scores. The loading bar was included to indicate movement and loading factors to indicate that the game is adapted to the player. The players have full transparency on how their scoring changes due to the game being adapted for a player by being able to see all factors. This aligns with requirement *R5: Make adaptations transparent*.

5.2 Dynamic Adaptation of Difficulty

The type of adaptation is explicit as players will be able to see a higher amount of bonus blocks spawned when they receive assistance in the game. The loading of the bar is multiplied by a factor that can increases, which leads to a faster fill of the loading bar. Consequently, the player who is lagging behind will unlock bonuses at an accelerated rate. This factor increases when the score difference between players increases and when more mistakes are made. Thus, the factor that fills up the loading of the bar faster, increases with the total amount of mistakes but also by the number of mistake streaks that are higher than 2. A mistake streak is the number of mistakes that were made in a row.

The adaptation can influence both players, as it initiates adjustments for the player who is currently lagging behind. The dynamic aspect of the adaptation and targeting both players fulfills requirement *R1: Create close games* as it should continuously balance the scores of both players. Furthermore, as both players are able to receive bonuses *R8: Both players need to get positive adaptations* is fulfilled. However, the

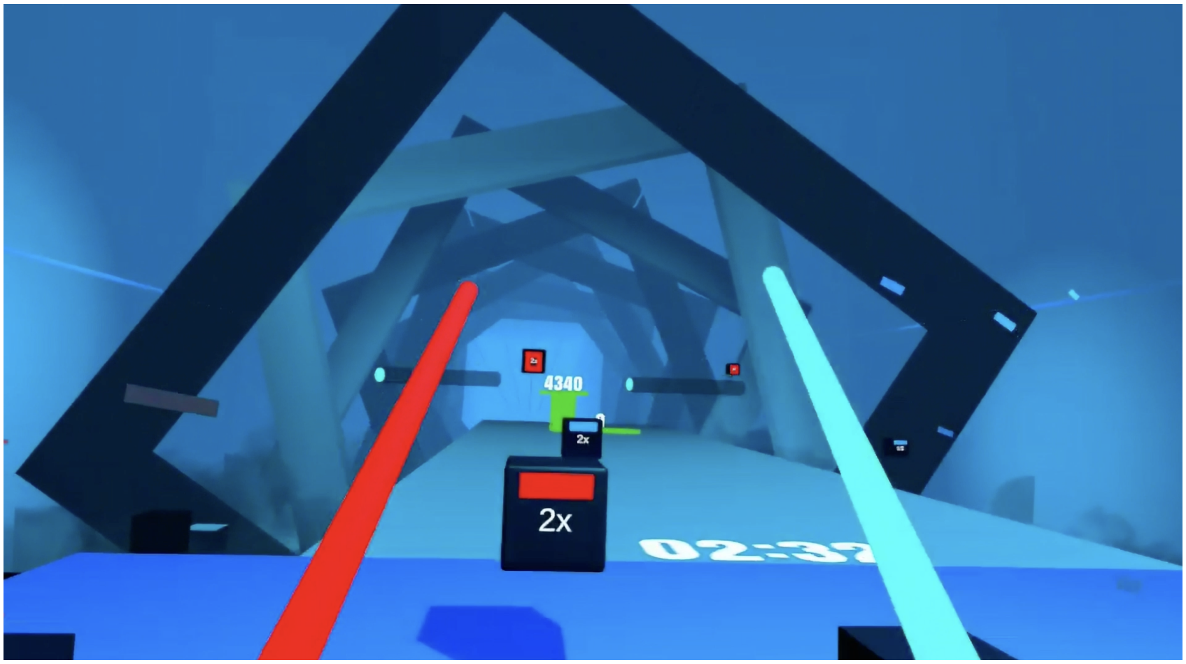


Figure 5: Exergame after bonus blocks activation

adaptation only comes into effect when the score difference exceeds 200 points, given that reaching 200 points is achievable by accurately hitting two blocks with high acceleration. Therefore, if the score difference between players is 200 or below, it is easy to secure a lead.

After playtesting the initial version of the game, an upper threshhold for the loading factor of the bar had to be added. Otherwise, there was the possibility that too many bonus blocks would spawn at the same time, which made the game impossible to play. After play testing, it was decided to set the threshhold to 28 as the amount of bonus blocks spawned at this loading factor was still manageable.

Additionally, the power-ups and attacks adapt depending on how the performances of both players differ. The first power-up allows the player to double the points when hitting blocks. Therefore, bonuses have a subtle effect initially and increase their effect further if necessary. Thus, subtle adaptation can be achieved according to *R6: Subtle adaptations*. However, it is necessary to increase the bonuses' effects to balance large score differences between players. Furthermore, if the player manages to hit another power-up block within 10 seconds, the factor increases. For example, instead of doubling, score will triple. The attacks decrease the size of the opponent's blocks and halves their points on a hit. If more attack blocks are hit within 10 seconds, the factor will be further halved. For example, 1/2 becomes 1/4 which devides the opponents score when hitting a block by 4.

Once the score difference is greater than 1000, the bonus mechanisms change. Thus, the game will become easier for the player. At this second stage, the power-up block

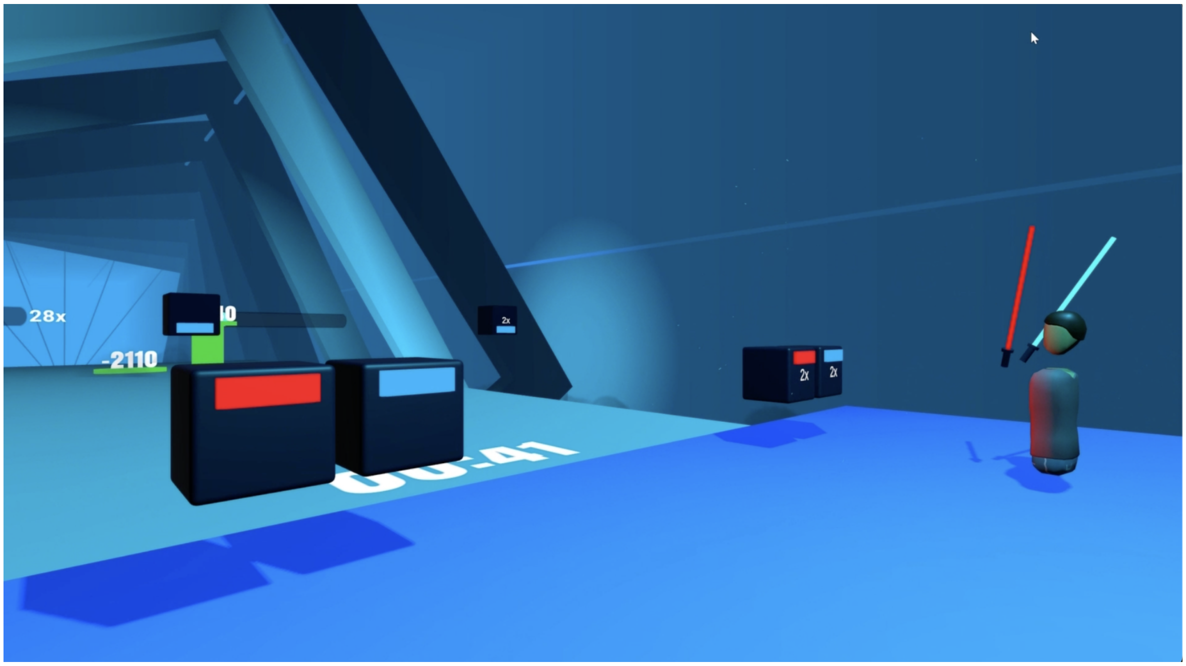


Figure 6: Opponent has activated power-up and receives double score (2x on blocks)

allows the player to hit the blocks in any direction but using the correct saber. This is indicated by colored bars on each side of the block. Red bars on blocks should be still hit by the red saber, while blue bars on blocks should be hit by the blue saber. When hitting the attack block, the opponent's blocks start to rotate which increases the difficulty. However, the opponent's score will not be decreased when hitting blocks.

The power-ups and attacks change again when the score difference becomes larger than 1500. At the third stage, power-ups allow to combine the previous two power-ups. Thus, when hitting the block, points will be doubled. Furthermore, blocks will not have any colored bars anymore and will appear completely black. This indicates that these blocks can be hit by any saber in any direction, which further simplifies the game. If the player attacks their opponent, the attack will be a combination of the previous two attacks. Therefore, the blocks' sizes will decrease, the points will be halved on a hit and the blocks will rotate. This further increases difficulty for the opponent.

5.3 Game Logs

For analysis certain values were logged during gameplay for each player, which were saved on the headset. Values were saved as a JSON file in key-value pairs and could be accessed by connecting the headsets via cable to a laptop. The overall score and a list of streaks were logged. The list of streaks included every streak that the player had during the game. For example, if the player hit 10 blocks correctly, a value of 1 was added to the list. If the player then proceeds to hit 10 more blocks correctly, it would add the value of 2. Additionally the overall amount of mistakes and a list of mistake

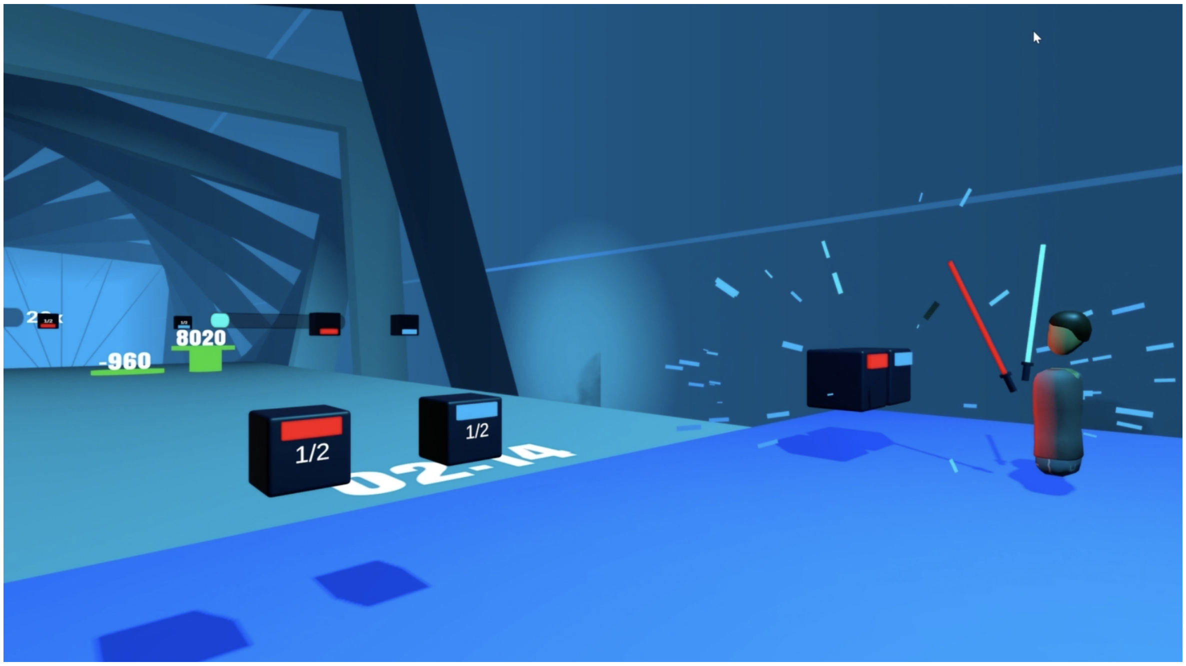


Figure 7: Opponent attacked player: 1/2 on blocks indicate half score

streaks were logged. A mistake was tracked when a player missed a block. If the player was missing continously blocks, the number would be counted and added to the list of mistake streaks. In case of the game with dynamic difficulty adaptation, it was also logged how the factor for the loading bar changed over time. Thus, each time a new factor was calculated, it was added to the list.

5.4 Play Testing

The prototype was play tested by four players, who were recruited through convenience sampling, including friends and work colleagues. Each player participated in 2 to 3 games with the researcher and was asked in between to describe liked aspects, potential improvements, understanding of game mechanics, and suggestions for improving the game. Following each play testing session, the game underwent adjustments based on feedback and was play tested again until a stable version was achieved.

The first play test highlighted the necessity of implementing a threshold for the adaptation factor, as a higher score difference continuously increased the factor, resulting in an excessive number of bonuses, which made the game unplayable. In response, adjustments were made to address this issue.

The second play test revealed User Experience (UX) problems that needed attention, such as excessive text and unclear scoring for correct hits, making it challenging for players to comprehend the scoring system. In order to make scoring clearer, score pops up from blocks when being sliced and disappear after 1 second. Additionally, it was unclear during the game who was currently leading. The play tester suggested incorporating a timer that emphasizes the urgency of gaining points before time runs out.

The third play test indicated that major issues were resolved, and players appreciated the UX improvements. However, catching up remained challenging. Therefore, it was decided to include three mechanisms for power-ups and attacks, which will further increase their effectiveness as the score difference increases. These changes made it easier to catch up even when score differences became very large.

After the fourth play test, no further suggestions were made, and catching up became achievable even with higher score differences.

5.5 Summary

The prototype for the VR multiplayer exergame was developed using Unity. The game mechanics, inspired by Beat Saber, involve slicing through blocks with sabers in different colors and directions. Additional features include bonuses that unlock with player movement, spawning two extra blocks, a power-up block to achieve higher scores, and an attack block to attack the opponent.

Dynamic adaptation of difficulty was implemented with a loading factor filling up a loading bar. This factor increases based on score differences and mistakes that players have made. When the loading bar is filled up, two bonus blocks will be spawned. These gives opportunity to catch up in the game by either making it easier to score higher or increasing the difficulty of the opponent's game.

Four play tests were conducted. Their feedback was implemented after each play test and evaluated in the subsequent play test. The first play test highlighted the need for a threshold in the adaptation factor. The second play test identified UX problems, leading to the incorporation of a timer for game progression. The third play test highlighted difficulties in catching up. It was decided to increase the power-ups and attacks effectiveness when the score difference increases. After the fourth play test, no further suggestions were provided, and catching up was feasible even with higher score differences. The prototype will be used in user experiments to further investigate how dynamic difficulty adaptation affects players.

6 Results Empirical Cycle

The following chapter describes the results derived from the analysis of survey data, logs, and interview responses. In evaluating the hypotheses, a quantitative analysis was conducted, providing insight into the potential rejection of the null hypothesis posited across the three aforementioned hypotheses. Furthermore, a combined quantitative and qualitative exploratory analysis has been undertaken to determine the influence of controlling variables on dependent variables and to identify potential advantages and improvements within the domain of dynamic difficulty adaptations in exergames.

6.1 Quantitative Analysis

For analyzing the survey data, a quantitative analysis was performed. Data were downloaded as a CSV file from Qualtrics with questionnaire results as numeric data. The results of the Likert scales were already converted to numeric values. All data processing and analysis of survey data and game logs have been done in R version 4.3.2. A mapping from statement to dedicated scores can be seen in Table 8. For reversed statements, scoring also has to be reversed. For example, a score of 7 should be reversed to 1. Thus, reverse statements in survey data received reversed scoring as a first step.

Statement	**Score**
Strongly disagree	1
Disagree	2
Somewhat disagree	3
Neither disagree nor agree	4
Somewhat agree	5
Agree	6
Strongly agree	7

Table 8: Likert scale scores

Afterwards, columns were added with further information such as the logged game data (scores, mistakes and adaptation loading factors), heart-rates and if the participant started the game with or without difficulty adaptation. Then scores for the scales competence, interest, effort and fairness were calculated. By multiplying the number of statements with 7 due to the usage of 7-Likert scales for all questionnaires, the maximum score can be determined. A lower score indicates less perceived competence, interest, effort or fairness, while a higher score indicates a higher degree of the perceived attribute.

Competence (PXI) has a maximum score of 21 as it consisted of 3 statements. Similarly, the maximum score of interest (IMI) is 7 * 6 statements = 42. The scale for effort (IMI) included 5 statements and therefore, has a maxmimum score of 35. Fairness has a maximum score of 63 as it included 9 statements.

Similarly, the maximum score for the social comparison score (INCOM) was calculated as 42 as it included 6 statements.

Furthermore, the physical activity (IPAQ) data had to be calculated. Participants stated the amount of days and minutes per week they had exercised vigorously, moderately, and walked. To determine how active they are, the amount of metabolic equivalent (MET) minutes per week had to be determined, which represent the energy used per minute to perform physical activity. One MET is the amount of energy used while resting. Walking is considered to be 3.3 METS, moderate activity 4 METS and rigorous activity 8 METS. Therefore, the participants' stated minutes per activity in the survey are multiplied by their corresponding number of METS [33]. The result is then recorded as MET minutes. Physical activities can be categorised into high, mid and low activity. Table 9 shows the necessary conditions that have to be fulfilled for each category. Each participant received one of the three categories to describe the extend of their weekly physical activity.

Part of the data processing was also to calculate the final score differences between the competing players for each game.

Outliers were not removed for any questionnaire answers or logged data as it is considered that answers can vary greatly as well as game logs for few participants.

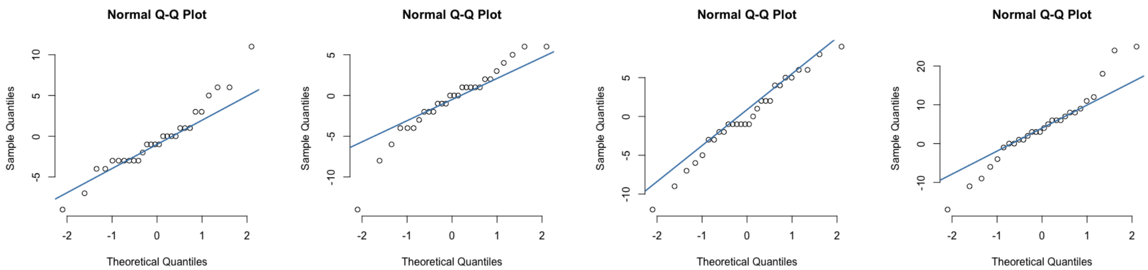
Category	**Conditions**
High	At least 3 days of vigorous activity + 1500 MET minutes a week OR 7 days of a combination of walking/moderate/vigorous activity + 3000 MET minutes a week
Mid	At least 3 days of vigorous activity OR 5 days of moderate activity OR 30 minutes walking per day OR 5 days of a combination of walking/moderate/vigorous activity + 600 MET minutes a week
Low	Not meeting criteria for mid or low activity

Table 9: IPAQ: Categories and conditions [33]

6.1.1 Hypothesis

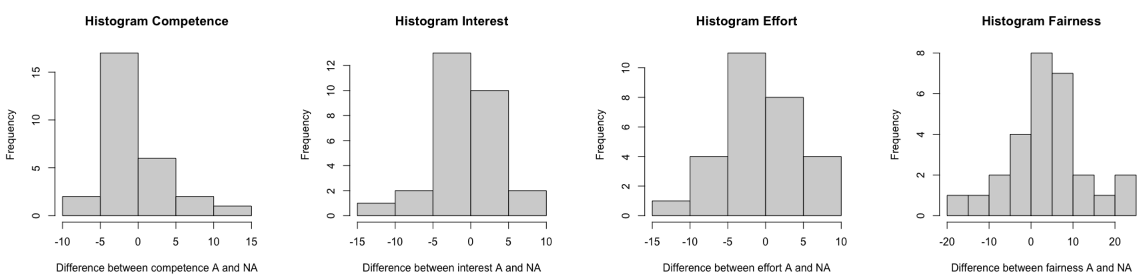
Following the preprocessing of survey data, the subsequent analysis was conducted using R. Preceding this, three hypotheses were formulated, positing a significant increase in the attributes of competence, motivation, and fairness in games featuring dynamic difficulty adaptation. To assess the rejection of the null hypothesis, a significance level of 0.05 was adopted and the cumulative scores of each scale for the game with adaptation were compared with the game without adaptation. For this, the parametric paired t-test was conducted. It was run as a one-tailed paired test to examine whether a game with adaptation would yield significantly greater values in any of the four attributes. These attributes corresponded to the four scales employed: competence, interest and effort, contributing to motivation, and fairness.

It is worth noting that opinions in previous research differ regarding whether Likert scales should be viewed as interval or ordinal data [9, 40]. While ordinal data allows



(a) Competence (b) Interest (c) Effort (d) Fairness

Figure 8: Q-Q-Plots of differences between a game with and without adaptation



(a) Competence (b) Interest (c) Effort (d) Fairness

Figure 9: Histograms of differences between a game with and without adaptation

ranking in order, interval data also satisfies the condition of having equal distances between data points. Likert himself regarded Likert scales as interval [9]. In this analysis, the sum of all values of a Likert scale was computed; hence, it can be treated as interval. To conduct paired t-tests, it had to be tested if the differences between the sums of a scale measured after each game were normally distributed. For all scales (competence, interest, effort, and fairness), the Shapiro-Wilk test was conducted to see if the distribution of data was significantly different from a normal distribution. For all four scales, the p-value was above 0.05, and therefore, it was normally distributed. QQ-plots were created and examined as well (Figure 8). Histograms in Figure 9 show the distribution but are more ambiguous due to the low amount of data points. Examining the Q-Q-Plots, histograms, and p-values of the Shapiro-Wilk test, it can be assumed that the data points are normally distributed for the differences between the scores of the adapted and non-adapted game in all four attributes: competence, interest, effort, and fairness. Therefore, a t-test was used to determine if there was a significant increase in any of those four attributes when adapting difficulty dynamically in exergames. The t-test was paired as a between-subject study was conducted. Furthermore, it was predicted that dynamic difficulty adaptation will only have a one directional effect by increasing the perceived competence, motivation and fairness. Therefore, the t-tests were one-tailed.

6.1.1.1 Competence

Hypothesis H_0: (H1.0) Dynamic difficulty adaptation in VR multiplayer exergames does not have an effect on players' perceived competence during gameplay.

Hypothesis H_1: (H1.1) Dynamic difficulty adaptation in VR multiplayer exergames increases players' perceived competence during gameplay.

Figure 10 shows that there was a higher range in competence scores for the non-adapted game, generally including higher and lower scores compared to the adapted game. The median of the non-adapted game was slightly higher as well with 16.43 compared to 16.04. The one-tailed, paired t-test showed no significant difference between adapted (M = 16.04, SD = 3.38) and non-adapted games (M = 16.43, SD = 3.8), t(27) = -0.5, p = 0.69. Thus, the null hypotheses H1.0 cannot be rejected.

6.1.1.2 Motivation

Hypothesis H_0: (H2.0) Dynamic difficulty adaptation in VR multiplayer exergames does not have an effect on players' motivation.

Hypothesis H_1: (H2.1) Dynamic difficulty adaptation in VR multiplayer exergames increases players' motivation.

Motivation was constructed by a combination of two scales, namely for interest and effort as high interest in the game yields higher motivation and higher effort indicates higher motivation as well. For both interest and effort it can be seen in Figure 11 that

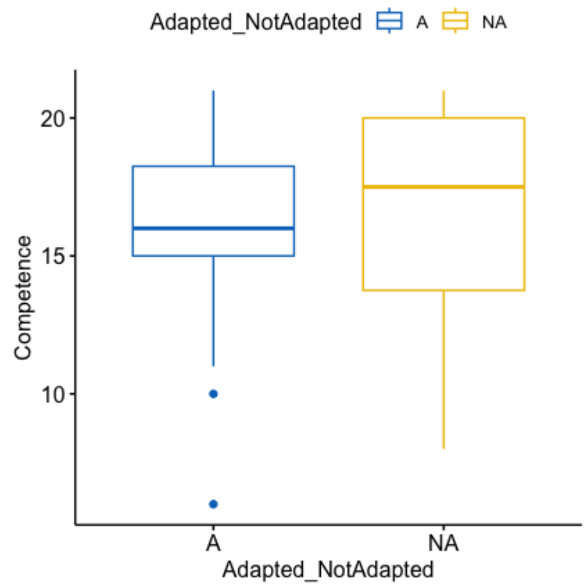


Figure 10: Box plot comparing competence between adapted and non-adapted game

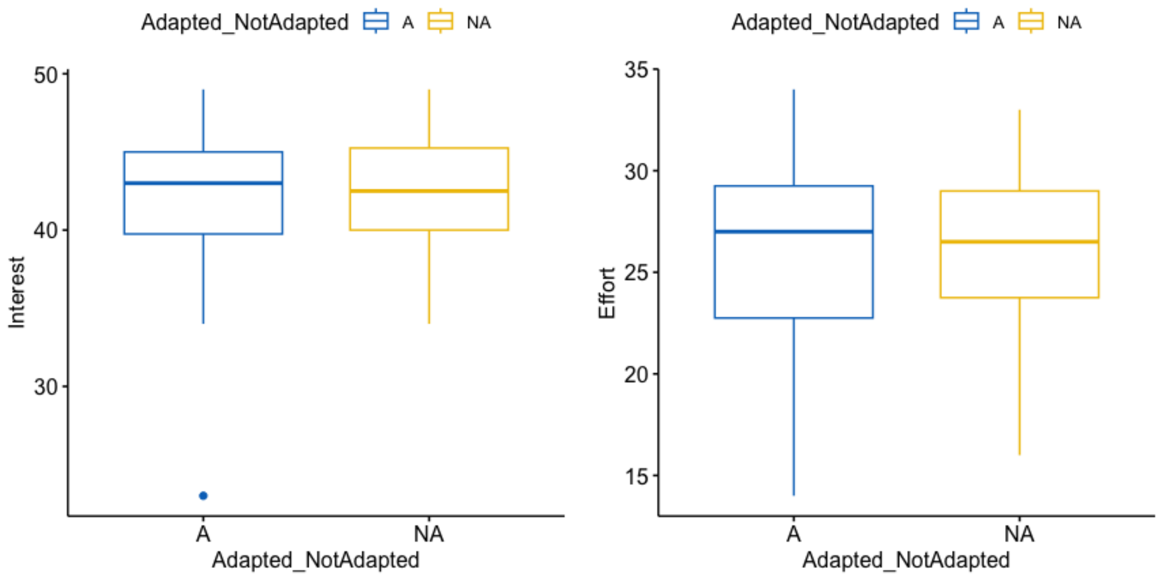
means are similar for both games. Furthermore, the ranges of scores between adapted and non-adapted game are similar for both attributes. Score ranges are above the mid-point of the score scale. A one-tailed, paired t-test showed no significant difference between adapted (M = 41.86 42.54, SD = 5.54) and non-adapted games (M = 42.54, SD = 4.07) for interest, t(27) = -0.84, p = 0.8. Similarly, there was no significant difference between adapted (M = 26, SD = 5.24) and non-adapted games (M = 26.04, SD = 4.53) for effort, t(27) = , p = 0.51. Thus, the null hypotheses H2.0 cannot be rejected.

6.1.1.3 Fairness

Hypothesis H_0: (H3.0) Dynamic difficulty adaptation in VR multiplayer exergames does not have an effect on players' perception of fairness of the game.

Hypothesis H_1: (H3.1) Dynamic difficulty adaptation in VR multiplayer exergames increases the players' perception of fairness of the game.

The box plot in Figure 12 shows that the adapted game has a higher median, while higher fairness scores were reached compared to the non-adapted game. For fairness, a one-tailed paired t-test showed that fairness was perceived significantly higher for adaped games (M = 47.89, SD = 7.86) compared to non-adapted games (M = 43.82, SD = 8.03), t(27) = 2.33, p = 0.01. Therefore, the null hypothesis H3.0 can be rejected. The user experiments have shown that the exact same exergame adjusted to have dynamic difficulty adaptation significantly increases the perception of fairness of players. The effect size was small with 0.44.



(a) Interest (b) Effort

Figure 11: Box plots indicating differences in motivation

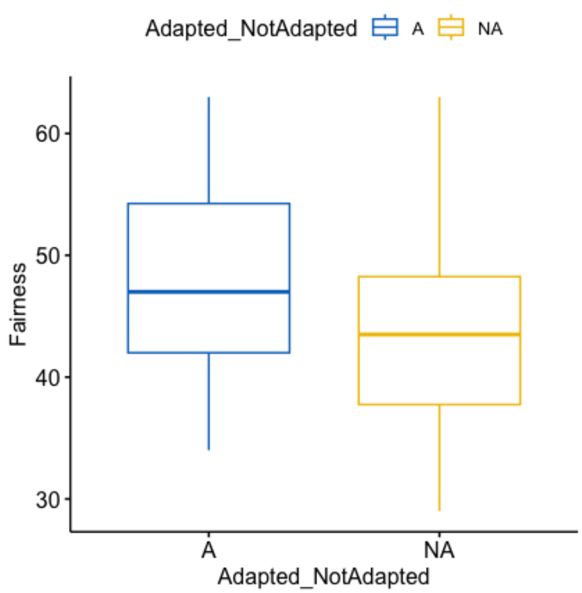


Figure 12: Box plot comparing fairness between adapted (A) and non-adapted (NA) game

6.1.2 Exploratory Results

Part of the thesis project is to explore whether other variables influenced the dependent variables. For this, an exploratory approach was taken. Weekly physical activity, heart-rates, the amount a person compares themselves to others (social comparison), gender and experience in rhythm dance games were recorded to explore possible influences.

6.1.2.1 Physical Activity

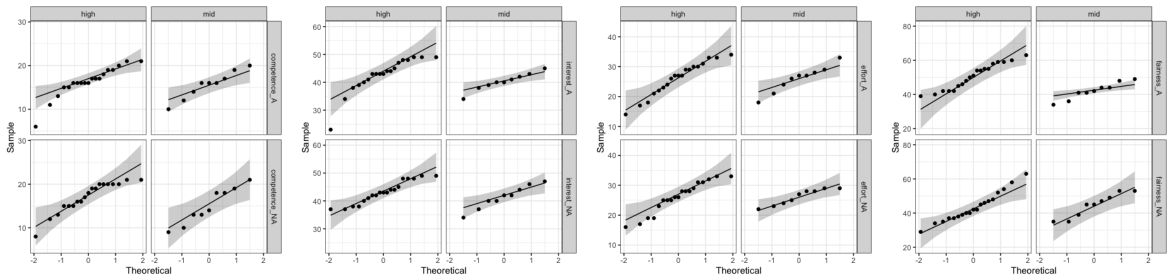
Most participants (20 or 71.5%) were highly active, while 6 (21.5%) were moderately active, and only two participants had low physical activity levels. Thus, the range of extents of physical activity was not varied enough to be able to test differences between different levels. Therefore, it was not possible to see whether the levels of physical activity influenced results.

6.1.2.2 Social Comparison

The survey entailed one question with 6 statements about social comparison using a 7 Likert scale and thus, the maximum score was 6*7=42. To facilitate comparison, the social comparison score was divided into three categories: low, mid and high. Dividing the score any further would have resulted in less participants per category. Similarly, dividing the score into two categories (high, low) would have resulted in 27 participants with a high score and only one with a low score. Thus, it was decided to determine three categories.

A scoring of above 28 was considered as a high score as it is over 2/3 of the maximum score. Respectively, a scoring above 14 was considered mid and any score equal or below 14 low. Note that 19 participants reached high social comparison scores, 9 had mid-level scores while no participant had a low score. Therefore, only mid and high social comparison scores can be analysed.

In order to determine whether there had been significant effects on dependent variables based on the level of social comparison, a mixed ANOVA was chosen with (non)existence of difficulty adaptation as a within-subject factor while the level of social comparison was considered as between-subject factor. The data was tested for normality with a Shapiro-Wilk test and by examining Q-Q-Plots (Figure 13). For the dependent variables effort and fairness, the Shapiro-Wilk test resulted in a p-value of over 0.05 (effort p=0.53, fairness p=0.29) and therefore the data is normally distributed. This can be also seen in the Q-Q-Plots (Figure 13). However, for competence the p-value was 0.005 and for interest 0.001. Even though the data points on the Q-Q-plots for both competence and interest (Figure 13) are mostly within the band, due to the small sample size the results of the Shapiro-Wilk test will be considered for determining normal distribution. Thus, the data for competence and interest is not normally distributed and therefore ANOVA could not be used. As a non-parametric alternative a Kruskal-Wallis test was conducted for competence and interest. However, only effects of the levels of social comparison could be investigated without factoring in the



(a) Competence (b) Interest (c) Effort (d) Fairness

Figure 13: Q-Q-Plots for adapted/non-adapted games and low/mid/high levels of social comparison

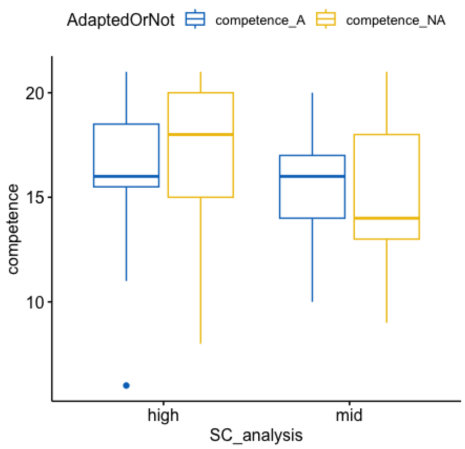
within-subject factor of adapted/non-adapted games. For example, it is not possible to see whether social comparison has effects only if difficulty is adapted. Furthermore, non-parametric tests yield less robust conclusions compared to parametric tests.

In order to conduct the mixed ANOVA test on effort and fairness, the data had to be tested for homogeneity with the Levene's test. ANOVA assumes that the variance between groups is equal the overall variance. For both effort and fairness, p-values were above 0.05 and therefore, homogeneity of variances was given.

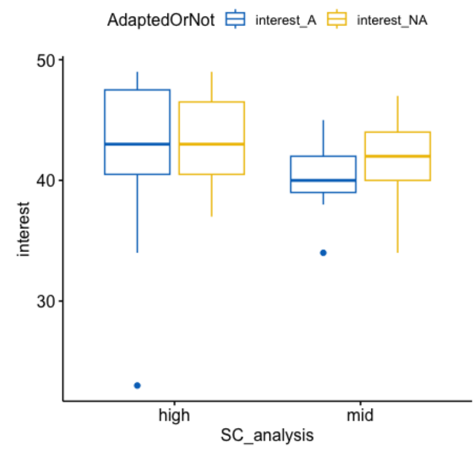
The results of the mixed ANOVA test showed that the simple main effect of social comparison on effort was not significant for adapted games ($F(1, 26) = 0.02$, $p = 0.90$).

A mixed ANOVA showed that social comparison has a significant effect on fairness, which differs between adapted and non-adapted games ($F(1, 26) = 8.38$, $p = 0.01$). Considering the Bonferroni adjusted p-value (p.adj), it can be seen that the simple main effect of social comparison on fairness was significant in adapted games ($F(1, 26) = 9.42$, $p = 0.01$). It was not significant for non-adapted games ($F(1, 26) = 0.11$, $p = 0.75$). It can be observed that, in pairwise comparisons, the mean fairness score was significantly different for high social comparison scores above 28 vs mid social comparison scores above 14 ($p = 0.01$). Participants comparing themselves more seem to value fairness higher for adapted exergames than other participants, which can also be seen in the box plot for fairness comparisons between high and mid-level social comparison (Figure 14d).

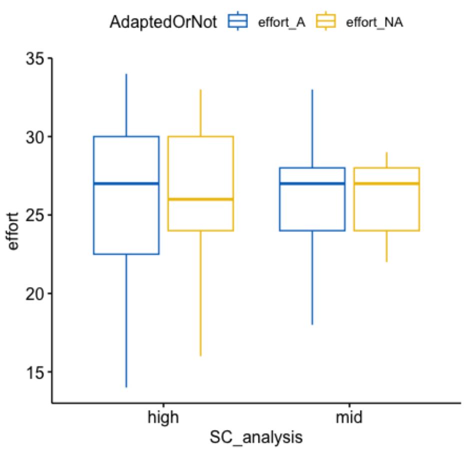
Conducting a Kruskal-Wallis test for competence showed that there is no significant difference for varying social comparison levels ($H(1) = 2.09$, $p = 0.15$). The Kruskal-Wallis test for interest showed that there is a significant difference for varying social comparison levels ($H(1) = 4.17$, $p = 0.02$). The effect size was calculated using eta-squared and revealed a small magnitude with an effect size of 0.06. As only two levels of social comparison existed in the data (mid and high), there is a significant difference in interest scores depending on whether participants considered comparing themselves



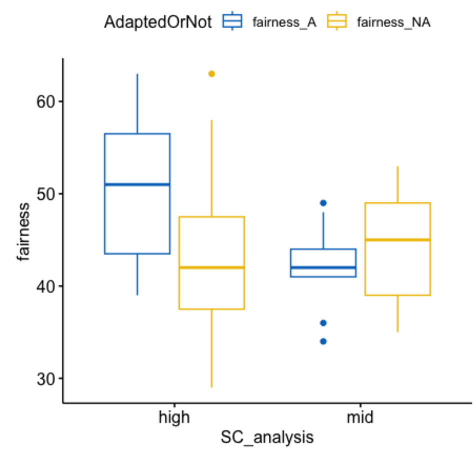
(a) Competence



(b) Interest



(c) Effort



(d) Fairness

Figure 14: Box plots for adapted (A)/non-adapted (NA) games and social comparison (SC)

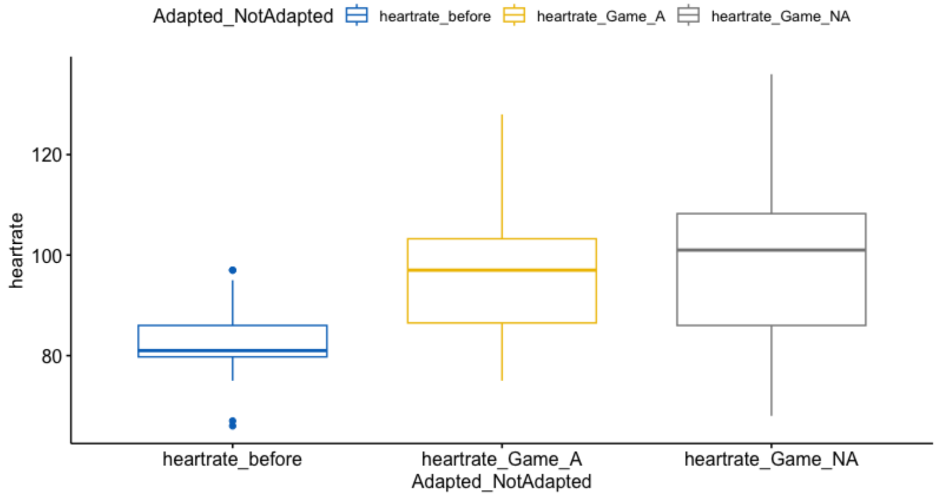


Figure 15: Box plot heart-rate in adapted (A) and non-adapted (NA) game

with others moderately or highly.

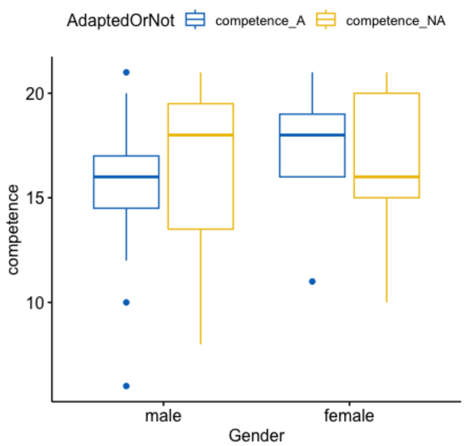
These results should be viewed carefully as only 6 participants had mid level scores.

6.1.2.3 Heart-Rate

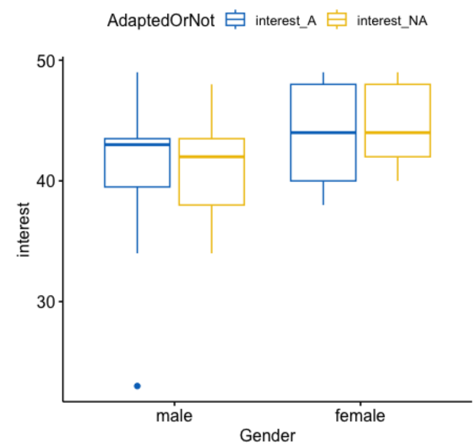
Elevated heart-rate during gameplay can indicate if players were physical active, put effort into playing and thus, were motivated to play. To determine if adapted difficulty had a significant effect on heart-rate compared to the non-adapted game, a Kruskal-Wallis test was conducted. A one-way ANOVA was excluded as Levene's test conducted for testing homogenuity of variance resulted in a p-value below 0.05. The Kruskal-Wallis test showed a significant difference between heart-rates ($H(2) = 19.8$, $p<0.001$). Furthermore, the effect size was large with 0.22. A pairwise comparison showed that heart-rate was significantly different before playing the exergame compared to after playing the adapted game ($p<0.001$) as well as after playing the non-adapted game ($p<0.001$). However, no significant difference was detected between adapted and non-adapted games. This can also be seen in the box plot (Figure 15) as the heart rate after playing the adapted and non-adapted games seemed similar.

6.1.2.4 Rhythm Dance Gaming Experience

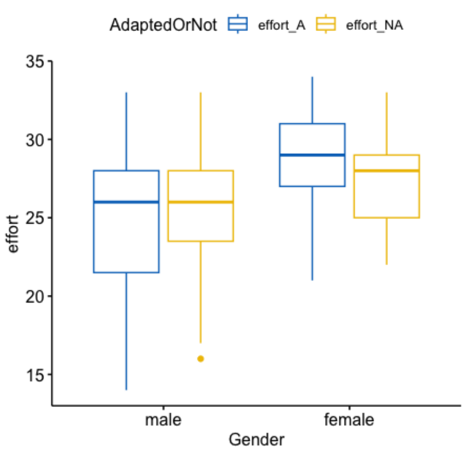
Only four (14.3%) participants indicated that they played rhythm exergames often or occasionally. More than half (19 or 67.9%) of participants had never played or had only played once or twice, and 5 (17.9%) had played rarely, meaning once every few months. Therefore, for most participants, rhythm dance games were novel. To detect effects, a broader range of experience is necessary.



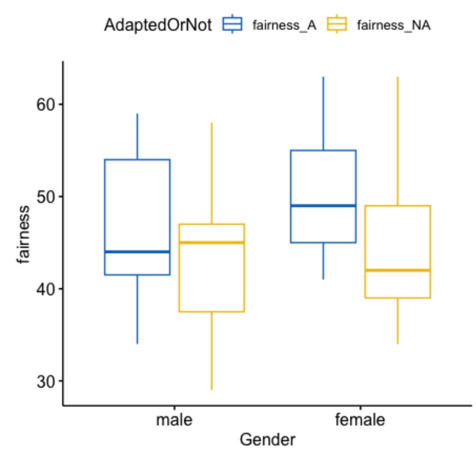
(a) Competence



(b) Interest



(c) Effort



(d) Fairness

Figure 16: Box plots for adapted (A)/non-adapted (NA) games and gender

6.1.2.5 Gender

The user experiments involved 19 (67.9%) male participants and 9 (32.1%) female participants. In line with the analysis of effects based on social comparison levels, mixed ANOVA tests for perceived effort and fairness were conducted, as all assumptions could be satisfied. Due to the non-normal distribution of data for competence and interest, the Kruskal-Wallis test was employed. As previously mentioned, this test can determine whether gender had a singular effect on the dependent variable. However, it does not provide insights into potential variations in effects between different genders and adapted or non-adapted games. The results should be carefully considered due to the low amounts of female participants.

The box plot in Figure 16a illustrates that female participants seem to perceive themselves as more competent when playing the adapted game compared to male participants. This reverses however for the non-adapted game. However, the Kruskal-Wallis test showed that there is no significant difference between male and female participants concerning their competence perception when not accounting for distinctions between adapted and non-adapted games ($H(1) = 1.36$, $p = 0.24$).

Concerning interest, there seems to be no apparent disparity between adapted and non-adapted games for participants of one gender, as depicted in Figure 16b. Overall, female participants seem to express higher interest in both games compared to their male counterparts when playing any of the two games. This difference was statistically significant according to a Kruskal-Wallis test ($H(1) = 5.70$, $p = 0.02$). It had a moderate effect size of 0.09.

Examining the box plot for effort in Figure 16c, it seems that more female participants put particularly more effort in the adapted games. A mixed ANOVA revealed no significant effect on effort when accounting for both genders and the adapted/non-adapted game conditions ($F(1, 26) = 0.56$, $p = 0.46$).

The fairness box plot (Figure 16d) suggests that, there were higher values selected for both genders and for the adapted game compared to the non-adapted game. No statistically significant difference could be detected according to a mixed ANOVA considering both gender and adapted/non-adapted games ($F(1, 26) = 0.44$, $p = 0.51$).

6.1.2.6 Player Scores

The objective of dynamic difficulty adaptation is to maintain game balance consistently, thereby creating an engaging and motivating experience for players. To assess the effectiveness of this adaptation, the disparities in player scores were compared between the adapted and non-adapted game. Employing the Shapiro-Wilk test and inspecting the Q-Q Plot (Figure 17), it was ensured that the score data was normally distributed ($p=0.06$). Hence, a paired, one-tailed T-Test showed a significant difference between adapted ($M = 3787$, $SD = 8746$) and non-adapted games ($M = 3946$, $SD = 4750$), $t(27) = 4.95$, $p<0.001$. The effect size was large with -0.917. The negative value indicates that the score difference decreased. This statistical analysis confirmed that

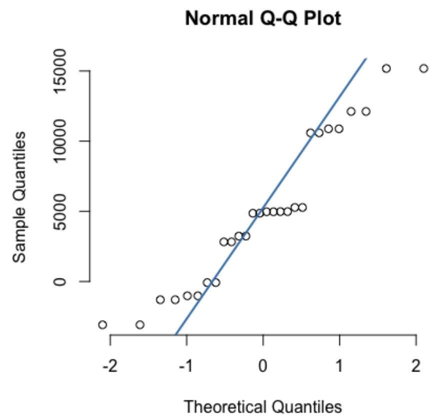


Figure 17: Q-Q-Plot to examine normal distribution in score data

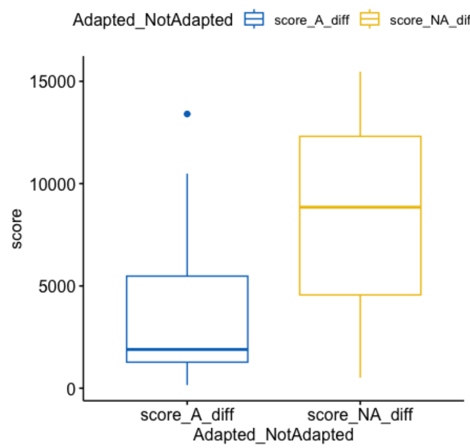


Figure 18: Box plot for score difference in adapted (A) and non-adapted (NA) games

the exergame featuring dynamic difficulty adaptation significantly reduced the score differences between the two players. Examining the box plot in Figure 18 it can be seen that the median is much smaller for the score difference in adapted games.

6.2 Qualitative Analysis

Part of the user experiment was a short interview conducted with both participants at the same time. This resulted in 14 recordings created on an Iphone with the app Voice Memos. All recordings were transcribed using the transcription automation software Sonix. Transcriptions were downloaded as pdf files and further analysed with the help of the software Nvivo version 14. For analysis emergent coding according to Straussian Grounded Theory was chosen. Thus, coding started with open coding by finding new codes or attaching statements to previously found codes.

Since participants were asked during the interview to state, which game they preferred as well as advantages and disadvantages of the dynamic difficulty adaptation implementation, they were added as codes as part of priori coding. Afterwards, codes were categorised into larger themes, which is called axial coding within the Straussian Grounded Theory. Hence, codes that are conceptually linked are categorised under one theme. As last part selective coding was conducted in order to find underlying theories based on the found categories.

Appendix E.1 shows a table with all found codes and categories. Codes have the number of found statements attached. Note that participants may state multiple times a certain statement and thus, the number of statements can exceed the number of participants. Furthermore, certain statements could be assigned to multiple codes. Seven categories were found through axial coding, which will be explained in the following sections.

6.2.1 Favourite Game

The majority of participants stated their favorite game either when choosing which game they liked more or when asked which game they would rather play again. Thus, 20 statements were found to support dynamic difficulty adaptation, while 8 participants stated that they preferred the game without adaptation. However, 6 participants of these 8 stated that they see value in dynamic difficulty adaptation and would prefer it in certain situations or see them as equally good.

> P1B: I guess I would say with assistance, I would say it was slightly more satisfying.

> P3B: But that's me as in a personal perspective, I like to be actually playing a game with your own, like, capabilities.

> P5A: Do I have to choose. For me they were very similar except that the first one had a higher challenge because of the resizing boxes.

6.2.2 Advantages

It was mentioned 28 times that the game including dynamic difficulty adaptation was fairly balanced. Thus, dynamic difficulty adaptation created the perception that every player had a chance in winning without overbalancing the game for one particular player.

> P2B: For me, it was nice because it motivated me more to catch up because the defense was too high, uh, in the first game [without adaptation].

Ten statements were found about the characteristic of dynamic difficulty adaptations to increase the challenge of skilled players. This has been seen positively, as skilled

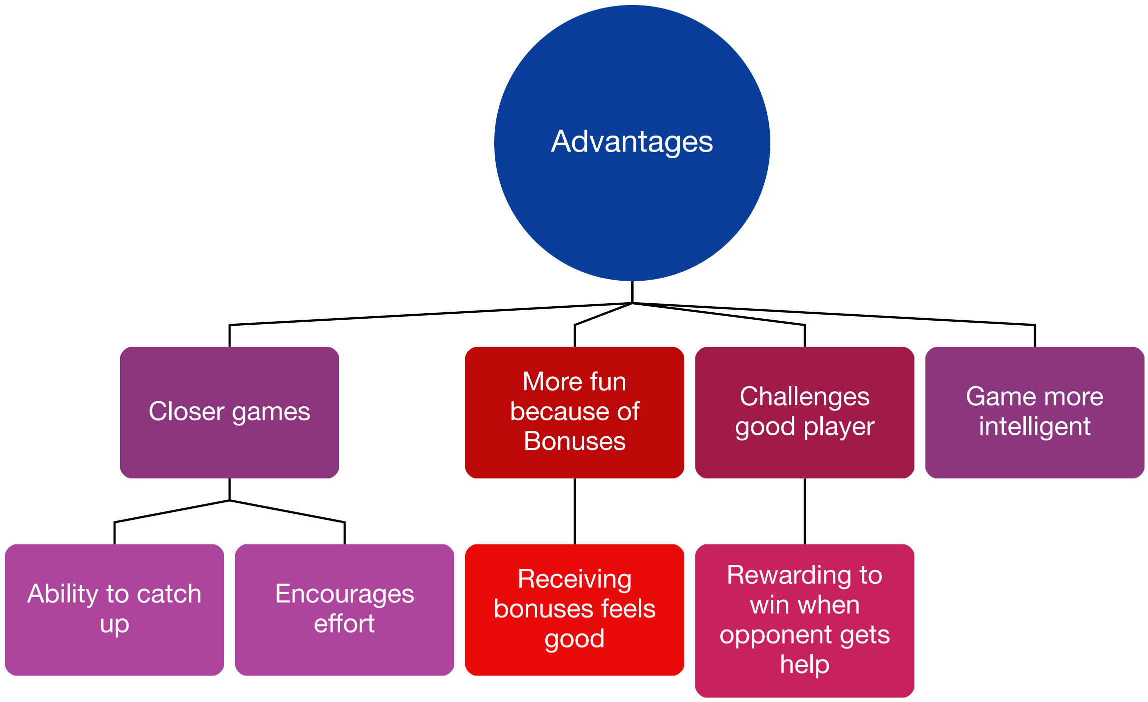


Figure 19: Mind-map: advantages of dynamic difficulty adaptation

players had to continuously put effort into playing to win. In the game without adaptation, score differences were becoming big enough at a certain point that skilled players were sure that they would win and, therefore, did not have to put as much effort into playing. Additionally, participants mentioned that winning the adapted game while the majority of bonuses received by the opponent felt rewarding.

> P11A: So like, yeah, at some point it's like, I'm gonna win. You're gonna lose. [...] Because it was really kind of neck-on-neck. You're gonna be like, oh, let me put in just a bit more effort.

> P3A: I feel like it makes it a little bit more tricky and makes you feel better when you get as much points with the adaptive thingies. So it feels like accomplishments. Okay, whatever you throw at me, that doesn't really matter. I'm good either way. So that was really nice.

This was also part of the reason why participants stated that dynamic difficulty adaptation increased their motivation. Nine statements were made about dynamic difficulty adaptation making the game more motivating to play. Another reason was stated that the adapted game was more dynamic, fast paced and therefore, more interesting to play. It was also perceived as more motivating as participants received more points through bonuses.

> P10B: At a certain point it became even boring a bit at the end because it was like, yeah, quite steady regular. The [adapted game] was much more fast paced, much more challenging.

> P8A: It was faster. There was more going on. So it had like a bit of a surprise element. And then also that I had to make a fast decision. Do I want to hit my color or take the bonus.

Furthermore, participants overall stated 5 times that they liked earning bonuses instead of getting them without any effort. Even though participants may have received a lot of bonus blocks, the fact that they were hard to reach made them feel like, they were not particularly helped in the game. It was also mentioned 5 times that receiving bonuses generally felt good. An overview of all mentions regarding advantages can be seen in Figure 19.

> P11B: I like that [the bonus] was kind of on the side, so I had to really work to get it and didn't always. I kept missing it a lot of times, so it didn't feel like I was really getting help, per sé. It was like, if you can get it, you can get a bonus, but it didn't feel like a handout or something. So I like that.

> P10A: It was quite satisfying to see, especially because I hit the stars a few times and I got to an eight times multiplier. And then I was like, oh, yeah, nice, nice, nice nice. So it felt satisfying to be able to like really bump up my points.

6.2.3 Disadvantages

While the challenging aspects of spawning bonus blocks have been seen positively, participants stated 21 times that the position of the bonus blocks made it too difficult to reach. Spawning bonuses also added more blocks to their game. Participants experienced especially in the adapted game that the constant flow of bonus blocks created a more challenging game as sometimes four blocks at once instead of one or two were spawning. Furthermore, participants had to consciously choose which bonus block to hit to either increase their own score or increase their opponent's game difficulty. This made the adapted game mentally more challenging during a fast paced game.

> P7B: I didn't feel like it got easier, I found I was behind and it got harder because there were more things to hit, and I was too busy trying to hit the bonuses to hit the actual blocks. Whereas I preferred the second game [without adaptation] because there was not a sudden bonus to try and hit, I could just focus on hitting the right color block with the right color saber.

> P10B: It was difficult sometimes to get them because I couldn't move too much from my spot, but the blocks were like pretty far away sometimes. So that was a bit difficult to reach.

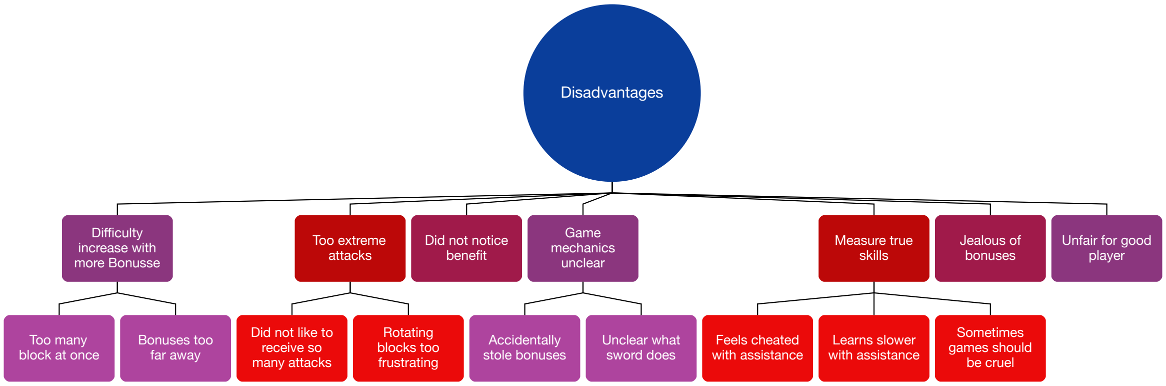


Figure 20: Mind-map: disadvantages of dynamic difficulty adaptation

Certain interface parts or game mechanics of the game were mentioned that were confusing for participants. These parts can be seen as problems in the user experience (UX) of the game such as how certain information was presented. Fifteen statements were collected about UX problems. Some participants did not understand what fully black blocks meant or that they have been attacked, which caused their blocks to suddenly change. It was also mentioned that the they might understand the game better after playing it more or that they only understood the game mechanics fully when playing the second game.

> P3A: I didn't really understand what did what before we began. So, I didn't know what happened to me once it happened. And maybe it could be confusing or not as well understood. I think that's the only thing I didn't really like at first. But if you play it more you'll learn what it does.

> P9A: Was more like the second one because the first one was more discovery part for me. I was like, okay, trying to figure out everything. And the second one I was like, okay, now I am ready and I was winning.

As a UX problem it was mentioned that they did not receive feedback when they hit the attack bonus block which had a sword as icon in the front of the block. Thus, participants hit this bonus block without understanding the benefit of it. In order to see the result of the attack bonus block, players would have to look at their opponent's blocks. However, participants focused more on their own blocks due to the high pace nature of the game.

> P6B: It was a bit unclear what the sword did, though. I only hit it once, and I didn't really get feedback from what would happen. I thought, okay, I hit the sword and I was happy that I hit the sword, but I didn't know if it was doing anything.

As mentioned before, the bonus blocks were spawning on the left and right side but further away from the other blocks. Some participants either mistakenly hit their opponent's blocks or thought that their bonus blocks were meant for their opponents. Thus, they did not attempt to hit these blocks. The position of the bonus blocks was therefore another UX problem.

> P13B: There was an issue to me, my standing position was a little bit too to the right, becausemwhen the stars come, I thought it was his block, so I didn't hit it. I didn't try to hit it at all. Only in the end, I find, oh, it should be mine. I should step a bit to the left.

The type of attacks were partly perceived too extreme as participants stated 15 times. Especially rotating blocks made it hard for participants to correctly hit the blocks. Some participants received constantly attacks in the adapted game due to their opponent hitting the attack bonus blocks often in a row. This can become frustrating.

> P13B: But at some point they also become a little bit frustrating in a sense that... So I can assume if you get that very often then the game becomes a little less enjoyable.

> P10A: I think it's specifically in the [adapted game] when like the blocks were spinning. I was like, I mean, I guess it is the whole point, of course, but it was like really confusing. I didn't really know where to hit. So I was just kind of just going for it. And perhaps that felt a bit frustrating, but I can appreciate that's probably the point.

Five times it was mentionend that either participants are more interested in comparing their true skills with other players or that it might become unfair to adjust games differently for each player. Figure 20 shows an overview of all mentioned disadvantages of dynamic difficulty adaptation.

> P14B: It somehow sounds unfair to the other player that now I'm not only playing against the player but also the AI is somehow helping or the game is helping me, but maybe it will increase the deficit, the competition.

> P9A: I would go without any help to have the fairness for each opponent and try to see where I am.

6.2.4 Improvements for Future Games

In total, 14 suggestions were made to improve the game (Figure 21). Three of them were only mentioned once, four were mentioned twice and one was mentioned 3 times. Thus, these suggestions should be carefully examined for their adherence to previously derived requirements through focus group discussions or the found theories through

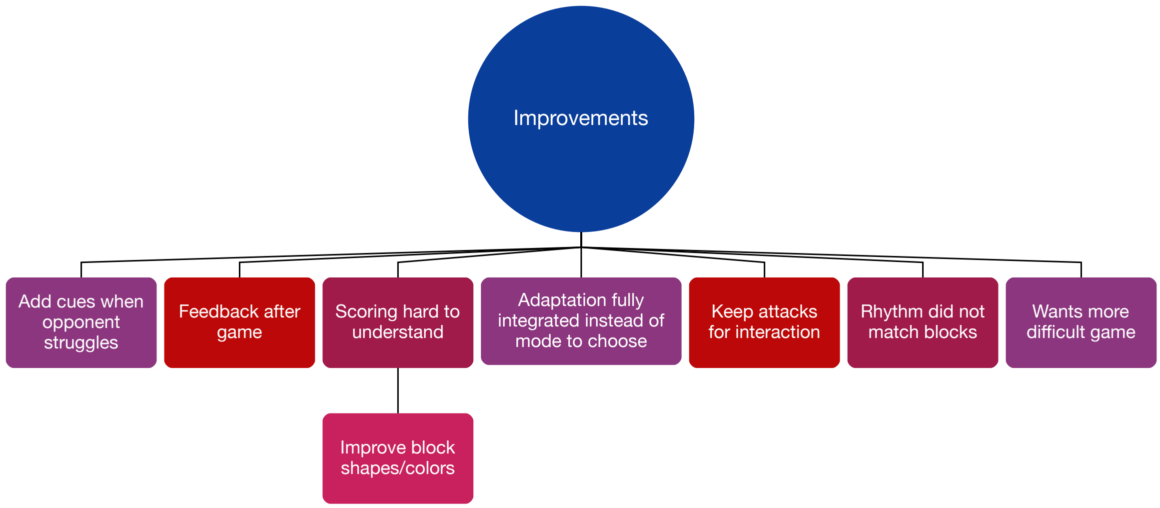


Figure 21: Mind-map: improvements for the game

the analysis of the interviews. Three statements were about the attacks and how they added interactivity in a rhythm dance game where players are usually not interacting with their opponents. Thus, attacks were positively seen. Note that these 3 statements were all made by one participant.

> P13A: The one thing I like is, is giving my opponent some disadvantage. Because there are some ways we can interact with each other. Yeah, that's missing in most of these kind of games like Beat Saber.

A participant pair, who played against each other, noted that they rather want dynamic difficulty adaptation to be automatically integrated as opposed to offer it as an additional mode.

> P12A: So if you choose assistance and you feel like you're... You know what I mean, right? Then you're choosing something which is giving somebody a handicap.

Another participant pair mentioned that adding feedback after the game to understand their current true skills and improvements in the game would be helpful. Hence, the feedback could include hints how to improve their game play.

> P9B: It's not like I don't like about it, but I would see it's more valuable that maybe the game can take it further that it gave me, after the game, like a little bit of tips that I should consider to actually learn, like why I was following that. What was better about the other opponent that I was missing? So by that I'm developing my skill. Otherwise on the long run I will stay in my place just with the help. I'm good without the help I'm losing.

The scoring when hitting blocks was confusing for one participant. For example, the participant might have thought that they hit a block correctly but did not receive a score. Furthermore, it was not clear how the scoring works in edge cases such as hitting blocks at a corner. This could be also improved by choosing different shapes for the colored shapes on blocks that indicate where to hit a block.

> P10B: And it was also not much clear to me if I was actually getting a box or not, because sometimes, like with the right saber, I was hitting the one of the other color and I don't know if it was counting points or not. It was not easy for me to understand if it was a good shot or not. You know, like, for example, when you do, uh, just dance or those games, you easily know or like Guitar Hero, you easily know. If you're hitting the points and if you're building the combo here. It was a bit difficult to understand if I was actually getting them or not. [...] I would like to try with different kind of shapes – where the colors are placed because right now it feels a bit weird. Like, for example, if I hit a box in the corner, would it get it or not?

One suggestion was about making it more visible that an opponent was hit by an attack as players had to closely look at the opponent's blocks to notice a difference. Thus, a visual cue when hitting an attack bonus block in the peripheral vision of the player could be added. Furthermore, a visual cue could be added when their opponents are struggling during an attack.

> P7A: Change the color of the other person's score. So when they're struggling, it turns orange or something like as soon as there's a bonus activated against them, there's an orange color or a flicker or something.

One participant would have enjoyed a faster game, and another one noticed that the blocks did not spawn according to the rhythm of the song.

6.2.5 Best Context for Dynamic Difficulty Adaptation

Dynamic difficulty adaptation is most effective in games where there are substantial differences in skill levels, as indicated by six statements. Consequently, it is particularly well suited for casual games that are played infrequently or when new players seek to familiarize themselves with the game and learn without being discouraged by frequent losses. An overview of all mentioned statement regarding context can be seen in Figure 22.

> P13A: So in case I play with someone that barely plays piano or barely uses VR or is very new to these things, then I would want to use the assistance option.

> P13B: But I try to introduce this game to a friend who never tried this kind of game. I will start with the [adapted] one, then it won't frustrate the person.

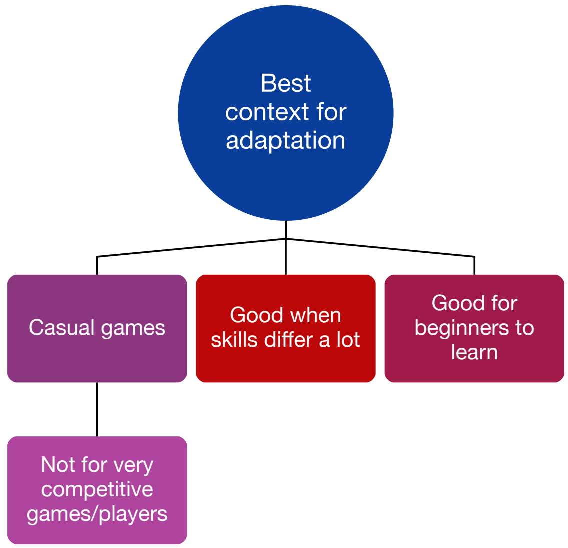


Figure 22: Mind-map: best context for dynamic difficulty adaptation

> P9B: I think at the beginning I will keep doing the one with the adaptive playing until a certain point when I learn a lot about the game, then I would go for the other type of games.

6.2.6 Other Notes

Some statements did not fit the other codes but were still interesting to look at. Two participants mentioned that they did not even notice that there was dynamic difficulty adaptation present as they focused on playing the game. One player mentioned that they were surprised to see that their opponent suddenly caught up.

> P8A: Yeah, I didn't really notice. I think I wasn't, like, consciously focusing on how much bonuses I was getting. I was just trying to hit everything that was going on.

> P3A: I feel like that the [game with adaptation] was a little bit weird because I was quite up for a long time. And then suddenly [...] the second player was almost on the same level, like, how is it possible? But it changed the challenges.

6.3 Derived Theories

Part of selective coding is to find connections between categories, which represent overarching themes within the interview data. To do this, the most common themes among the responses of the participants were taken into account. Four theories resulted from selective coding, which are explained in more detail in the following sections. Note that

UX problems were not included as part of the theories because they likely resulted from prototype shortages due to time restrictions in this thesis.

6.3.1 Motivate by Design of Dynamic Difficulty Adaptation

Participants mentioned multiple reasons why playing the adapted game was more motivating. First, the nature of close games created a challenge for skilled players. At the same time, less skilled players were encouraged to play with great effort until the end of the game because they felt they had the chance to win. Both players, regardless of skill levels, were motivated to put effort into playing until the end. Therefore, the design of dynamic difficulty adaption should first focus on how to achieve a close game. Furthermore, the particular implementation of dynamic difficulty adaptation further increased motivation. Achieving higher scores in a shorter time through bonuses, the added interactivity through attacking opponents as well as the high pace of the game, made the game play more interesting and satisfying. Thus, games incorporating dynamic difficulty adaptation should further enhance the game by adding new and interesting game mechanics.

6.3.2 Increase Difficulty Carefully

Dynamic difficulty adaptation can be implemented bidirectionally, involving both increasing and decreasing difficulty. The prototype developed for this thesis project facilitated bidirectional adaptation, leading to an increase in difficulty for skilled participants. Consequently, the scores were consistently balanced throughout the game, presenting a challenge for skilled players to maintain a lead. This challenge was viewed positively, as it motivated players to invest effort continuously in the game. However, it is essential to maintain a balance between challenging and overwhelming players, as outlined by Flow theory [25]. Therefore, adapting difficulty levels must be carefully calibrated to ensure that players perceive the increased difficulty as conquerable with additional effort or increased skills. For example, in this specific prototype adding a way to shorten attacks or lessen the impact by certain actions or movements may decrease the overwhelm that participants experienced partially during the user experiments.

6.3.3 Players Earn Bonuses

Previous research found that participants may feel inferior if they receive explicit assistance in playing as they feel exposed for their inferior skills [35, 41]. Dynamic difficulty adaptation was implemented as an explicit approach by spawning a higher amount of bonus blocks. Thus, participants were aware that they receive assistance as they saw a higher amount of bonus blocks. However, the bonus blocks represent a possibility to catch up. This means that participants had to successfully hit the bonus blocks, which were located further away, without missing regular blocks. This increased the difficulty of the game, as participants stated that the game was more fast paced or dynamic. Participants had to put effort into reaching those blocks, and thus felt that they did not receive an advantage over their opponent. Even with more opportunities to catch up, it was still a matter of skill and strategy to be able to catch up. This was considered positive, as participants felt they could still catch up if they put more effort into playing, but did not think that they received scores for free.

6.3.4 Preference for True Skills

Participants who expressed a preference for the non-adapted game cited their desire to compare true skills as their primary reason. Consequently, balancing exergames with dynamic difficulty adaptation may not be universally applicable, depending on individual preferences and specific gaming contexts. Additionally, participants highlighted that dynamic difficulty adaptation is more suitable for casual games, which are played occasionally and for short durations. One participant suggested that dynamic difficulty adaptation could serve as an initial aid for getting accustomed to a game, gradually phased out as players become more confident or skilled. It is crucial to consider that dynamic difficulty adaptation in exergames is most effective in scenarios involving beginner players or significant skill disparities due to varying levels of experience.

6.4 Summary

The quantitative analysis indicated that players rated fairness significantly higher for the exergame, which included dynamic difficulty adaptation. However, concerning competence and motivation, there was no significant difference between the exergame with and without adaptation. The developed exergame aimed to achieve a balanced score difference between players, and the analysis demonstrated its success. Score differences were significantly lower for the adapted game, indicating that scores between players were more balanced. Additionally, it was observed that heart rate increased during gameplay, as measured heart rates were significantly higher after each game compared to before game play. However, no significant difference in heart rate was observed between the two games. Overall, female participants expressed significantly higher interest in the game compared to male participants. Moreover, participants with high social comparison scores perceived fairness and effort significantly differently compared to participants with mid-range social comparison scores.

Overall, the qualitative analysis of the interviews revealed that the majority of participants either expressed a preference for or recognized the value of the exergame with dynamic difficulty adaptation. Participants reported heightened motivation during gameplay, experiencing a continuous challenge while still perceiving the possibility of achieving first place. Criticisms were directed at the increased difficulty resulting from a higher number of blocks in the adapted version. Particularly, skilled players were impacted who faced more frequent and intense attacks, leading to frustration. This emphasizes the need for careful design of the difficulty increase. Some participants opposed dynamic difficulty, expressing a preference for comparing their true skills with others. Nevertheless, dynamic difficulty adaptation in exergames proves beneficial in scenarios with substantial skill disparities, especially during the learning phase for beginner players.

7 Discussion

Exergames have the potential to motivate people to engage in physical activities, even with busy schedules, by integrating computer games with sports. Playing with friends can further enhance the overall gaming experience. This thesis project specifically focused on rhythm dance games in virtual reality (VR) and developing an approach to dynamically adjust difficulty levels based on players' in-game performances. User experiments were conducted to assess the impact of dynamic difficulty adaptation on participants' perceived competence, motivation, and fairness. The results will be discussed in the subsequent sections in order to answer the research question:

How can dynamic difficulty adaptation be used to balance mismatching skills of multiple players?

The question is answered in two parts. In the first part, requirements were investigated as part of the design cycle. These requirements can be used as design guidelines for implementing dynamic difficulty adaptation in an exergame. The purpose of the requirements is to ensure the acceptance of dynamic difficulty adaptation and high motivation of players. The second part includes an empirical cycle, in which the implementation according to the requirements is evaluated and effects on players are investigated. The following sections discuss whether the requirements were useful and the effects found in the user study.

7.1 Comparison of Focus Group and Interview Results

In order to investigate how dynamic difficulty adaptation can be used in exergames, design requirements were derived from the results of the focus groups. However, designs created by participants and their reasoning behind their design choices might not hold when testing the designs in user studies. Thus, the following requirements derived from focus groups were compared with the user study results.

7.1.1 R1: Create close games

The prototype has successfully created a closer game. In interviews, participants mentioned that they noticed how the score difference was smaller in the adapted game. It was also mentioned that the adapted game motivated both skilled and less skilled players to put more effort into playing. The skilled player was motivated to maintain their leading position, and the less skilled player felt that they still had a chance to win. It can be concluded that when designing dynamic difficulty adaptation into exergames, the focus should be on creating a close game, which also mitigates overbalancing.

7.1.2 R2: Avoid demotivation of the less skilled player by player control

Prior research has found that balancing exergames with visually explicit approaches can lead to negative feelings in less skilled players as they might feel exposed as weak player. Even though this research chose a visual approach to adapt difficulty by spawning more bonus blocks, it did not cause such negative effects on players. The fast pace of the game made it difficult to focus on opponents. Furthermore, the adaptation only increased opportunities to catch up but did not decrease score differences without the player's input. Spawning bonus blocks also increased the game difficulty. Thus, less skilled players did not feel less competent even with an explicit balancing approach. This could be achieved by offering opportunities that can help to catch up, which requires effort. This aligns with the requirement to give the player control over the adaptation as players still have to actively choose to use a bonus.

7.1.3 R3: Make improvement visible

In both interviews and focus group discussions, participants expressed the desire to monitor their true skills. As a suggestion, it was proposed that players could have access to information about their true performance and achievements after the game. However, this feature was not implemented in this research, as the primary focus was on examining the effects of dynamic difficulty adaptation during gameplay. However, providing players with information on their achievements over time could potentially serve as a motivational tool, encouraging them to play and practice more.

7.1.4 R4: Add the right amount of randomness

It was discussed in focus groups that games should not be too random, as this would result in a game outcome determined solely by luck. However, randomness can enhance games by making them more unpredictable, and consequently more interesting. The prototype in this study incorporated minimal randomness; for instance, the spawn position of bonus blocks was randomly assigned. Additionally, the use of rotating blocks as an attack introduced an element of randomness, as players did not know where to hit the blocks, resulting in random hits. Given that this type of attack received criticism for its random nature, it is crucial to carefully integrate randomness into the game. The aim is to avoid creating an impression for players that the game's outcome is dependent on luck.

7.1.5 R5: Make adaptation transparent

The game was designed to make the adaptation as transparent as possible, incorporating loading bars that visually filled up with increased movement and a loading factor displayed in text, indicating the speed at which the loading bar filled. Due to the fast-paced nature of the game, participants found it challenging to comprehend all presented game elements. As a result, their focus was primarily on the blocks, with some participants not even noticing the score bars that indicate who was currently leading in the game. It can be concluded that for exergames, which are mostly fast paced, information must be carefully integrated into the game, prioritizing details that are important during gameplay such as score. Achieving complete transparency in the adaptation approach may be challenging. However, including this information after the game could be a potential solution. This could be combined with showing player achievements after the game.

7.1.6 R6: Subtle adaptation

The implementation of dynamic difficulty adaptation in this research included a method to enhance the effectiveness of bonus blocks when score differences increased. Therefore, it was possible to balance games between players with vastly differing skills. Interview results revealed that participants experienced frustration due to attacks being too extreme at times. This was particularly evident in the adapted game, where attacks and bonuses became more effective when score differences increased. Specifically, the attack that led to rotating blocks was frustrating as it became challenging successfully hit the blocks. Therefore, a balance must be found between adapting games of players with major skill differences and implementing game mechanics that do not overwhelm players, especially if the adaptation increases the game's difficulty for the skilled player. Furthermore, participants stated that they do not want to receive obvious help. Hence, they suggested a subtle approach. However, dynamic difficulty adaptation only offered opportunities to players that could only be used by increasing effort, was was positively seen. Thus, a subtle adaptation might not be needed. Instead, players should feel as if their performance was a result of their own effort and not the game's assistance.

7.1.7 R7: Seeing opponents effort through avatar movements

In rhythm dance games, each player focuses on their own blocks or movements. Due to the fast-paced nature of these games, players may not notice their opponent, even if positioned next to them. One participant stated in the interview that they did not observe the effects of attacks and expressed a desire for a visible indicator during gameplay. Consequently, it is unclear whether players would prefer having their opponent as an avatar next to them. However, participants appreciated the ability to attack their opponent, enhancing interactions between players in a game that typically lacks such interactions. Incorporating visual indicators, especially after an attack, could address this issue. These visual indicators need to be placed in the players' peripheral vision in order to be easily seen.

7.1.8 R8: Both players need to get positive adaptations

The dynamic difficulty adaptation approach was implemented to accommodate both players, regardless of their initial performance. Consequently, any player who falls behind has the opportunity to receive assistance. This likely contributed to the higher perceived fairness in adapted games, as no player is inherently advantaged. Therefore, when designing dynamic difficulty adaptation, it should be designed to address the needs of all players, depending on who is currently lagging behind.

7.2 Interpretation of Results

The following sections include the interpretation of the quantitative results, which are then compared to the results of previous research, the findings of the qualitative analysis, and the outcomes of previous focus groups. In order to understand how dynamic difficulty adaptation can be used in exergames, its effects need to be investigated. Hence, the first three sections specifically discuss the perceived fairness, motivation and competence of participants. Further sections discuss other effects that have been seen in the user study.

7.2.1 Dynamic Difficulty Adaptation Created a Fair Game

The quantitative analysis results reveal a significantly lower score difference in adapted games, indicating the successful balancing effect of the implemented dynamic difficulty adaptation approach in the exergame. By implementing an approach that allows to adjust difficulty dynamically and for both players, overbalancing could be mitigated. Prior research about balancing exergames have suffered from problems due to overpowering the weaker player [4, 35, 77]. Similarly to this research, Jensen and Grønbæ could mitigate overbalancing by incorporating a dynamic balancing method for both players [41].

Additionally, fairness was perceived significantly higher in the adapted game, possibly due to the successful creation of a close game. Previous research has suggested that close games contribute to a more exciting experience [27, 87]. Furthermore, participants likely found the adapted game more fair due to the notion that less skilled players have the opportunity to win by investing more effort, while skilled players remain continuously challenged. It is crucial to emphasize that both players maintain the belief that their success is a result of individual efforts. Therefore, even when a game provides assistance, it should aim to motivate players to exert more effort rather than simply awarding scores without the player's input.

7.2.2 Differing Results for Motivation

Analyzing the questionnaires revealed that participants did not feel significantly more or less motivated in the adapted game. Upon closer examination of the box plots for interest (Figure 11a) and effort (Figure 11b), the scores appeared similar for both attributes. It is important to note that the majority of participants were new to rhythm dance games. Furthermore, the prototype consisted of an exergame that entailed novel game mechanics such as the ability to influence the opponent's game. The novelty effect likely influenced the motivation of the participants. As a result, the scores for interest was fairly high, with the median and values above the mid-point of the score scale. The median of effort was above the mid-point of the scale. Similarly, most values were above mid-point as well. It can also be observed that there is a significant increase in heart-rate after playing the exergame compared to before the experiment started. Thus, participants felt motivated to play the game, putting physical effort into playing. Note that participants received instructions on how to play the game, where movement was emphasized. Thus, the game encourages movement, which could be another reason for the values in effort.

However, during interviews, participants stated multiple times that the adapted game created a more motivating experience by creating a closer game. Participants felt that they had opportunities to win until the end and were challenged more by their opponent. Therefore, it remains unclear whether dynamic difficulty adaptation may lead to higher player motivation when the novelty effect wears off.

7.2.3 Effects on Competence

No significant effect was found on perceived competence. However, the box plot (Figure 10) shows a lower median for adapted games and therefore, more participants selecting lower scores for competence. Some participants reported feeling less competent in the adapted game, as indicated in their interviews, citing the perception of increased speed and difficulty. While this had a negative impact on some participants, leading them to initially prefer the non-adapted game, others enjoyed the dynamics of the adapted game and finding it more exciting. Furthermore, the type of attacks was perceived too extreme, particularly in the adapted game, where attacks intensified when scores differed greatly and players had to endure numerous attacks. Therefore, difficulty must be carefully adapted to avoid overwhelming players while still creating a challenging and engaging game. This aligns with Flow theory, where players can experience either boredom or overwhelm if their game experience is not meticulously designed to keep them challenged but not overwhelmed [25].

7.2.4 Designing Explicit Balancing

Gerling et al. stated that explicit approaches for balancing could diminish a player's competence [35]. This phenomenon was also partly observed in the user experiments conducted by Jensen and Grønbæk [41], particularly affecting players with lower skill levels. Similar to the prototype developed in this research, Jensen and Grønbæk investigated dynamic difficulty adaptation. However, despite the explicit balancing approach developed in this research, there was no significant difference in participants' competence compared to the exergame without dynamic difficulty adaptation. Additionally, participants did not feel that they were at an advantage due to receiving more bonuses, as they still had to earn their scores. Instead, the explicit balancing approach increased the game's difficulty by spawning more blocks that needed to be hit simultaneously. Participants needed to make strategic decisions about which bonus block to hit in order to maximize their scores. One of the derived requirements based on the focus group results was to allow players to earn their score even if they receive help. Participants emphasized that simply giving away advantages for free would lead to an unfair game.

Furthermore, the extent to which difficulty adaptation is noticeable in gameplay may also play a role in its impact on the player's competence. For instance, Jensen and Grønbæk explored a physical balancing approach in which players had to move closer to the screen if they performed worse than their opponent. Since participants felt that they were put on the spot, it affected their sense of self-efficacy negatively [41]. In this research, participants were primarily focused on their blocks and did not have the time to monitor their opponent due to the fast pace of the game. They likely did not feel as if they were put on the spot, as they did not sense being observed by their opponent. Previous research has also found that people can feel insecure while doing physical activities in front of peers [62]. By exergaming in VR and focusing on their own game this might have mitigated the problem of declining self-efficacy. The fact that players are fully immersed in virtual reality may be the reason that people fear less negative reactions from peers.

Therefore, visually explicit balancing approaches for dynamic difficulty adaptation are acceptable, provided they are designed as opportunities that must be actively used by players. This implies that players need to activate bonuses by leveraging their skills, thus earning their bonuses. The decision to adopt a dynamic approach to adapt difficulty was rooted in the notion that overbalancing should be avoided. Hence, the implemented approach of dynamic difficulty adaptation, which allows players to earn their bonuses, proved effective in mitigating overbalancing.

7.2.5 Balancing is Unsuitable for Highly Competitive Players

The results of the focus groups and interviews reveal that some individuals may not accept dynamic difficulty adaptation in games. They expressed concerns about fairness, as competitive games are expected to assess skills. Interestingly, the quantitative analysis indicates that fairness was perceived to be higher in the adapted game. One possible explanation for this is that participants were not fully aware of which game they were playing.

Moreover, despite the visually explicit adaptation, participants did not feel as if they were given assistance during the game, as the bonus blocks also increased the game's difficulty. Another factor contributing to the acceptance of dynamic difficulty adaptation may be a person's level of competitiveness, which was not explicitly measured in the user experiments. Instead, social comparison was measured, and most participants achieved at least 2/3 of the maximum score.

Social comparison and competitiveness might be correlated, as competing with others involves comparing skills, but they are not synonymous. Individuals who compare themselves with others may prefer difficulty adaptation in games as it helps them win. However, competitive players may prefer to know how they stack up against others. Significant differences in fairness were observed between participants with high social comparison scores and mid-level scores for adapted games. A reason for this result could be that participants who compare themselves with others a lot prefer to receive assistance during the game in order not to lose. Furthermore, a significant difference in interest could be observed between participants of mid and high level social comparison without taking into account the factor of adapted and non-adapted games. Participants, who compare themselves more, may be more interested in competitive games in general.

7.2.6 Effects based on Gender

Mostly, there have not been significant effects measured between female and male participants. However, female participants expressed significantly higher interest in playing this exergame regardless of the presence of difficulty adaptation. Differences in gaming preferences have been researched before, and results of user experiments show that female players prefer less complex games in comparison to male players [86]. Simple games include rhythm dance games due to their intuitive gaming control through movement. It was also found that women tend to like music/dance games more than men [65]. Therefore, female participants may have included high scores for interest in both games because rhythm dance games are generally interesting for female players.

7.3 Limitations

The results of this study may have been limited by certain factors. For example, convenience sampling did not result in completely heterogeneous groups, and certain biases may have influenced the participants' responses. The following sections further elaborate on the limitations of the study.

7.3.1 Participant Sample

There were a total of 28 participants. However, it is important to note that certain effects may have arisen due to the combination of various variables. For instance, the acceptance of dynamic difficulty adaptation in exergames might vary among players with different experience levels in rhythm dance games. Ideally, each group with a specific experience level would require 30 participants to produce reliable results. Similarly, multiple variables, such as gender, the level of weekly physical activity, and the degree of competitiveness or social comparison, could influence the dependent variable. To adequately address these factors, multiple groups of 30 participants each would be necessary.

Statements made during focus groups and interviews suggested that highly competitive players might not readily accept dynamic difficulty adaptation in games, and it could be more beneficial for casual gamers or beginners. However, participants were not specifically recruited based on competitiveness or game experience. Most participants in the study had a high level of weekly physical activity. Additionally, most participants were relatively new to dance rhythm games. Consequently, no quantitative analysis could be conducted to substantiate these statements.

7.3.2 Experiment Context

The experiment took place in a single room for a participant pair playing against each other. It is essential to acknowledge that the gaming experience might differ if participants were playing remotely, where they could no longer hear each other in the same room, resembling a more realistic VR multiplayer game setting.

Each pair played two games, one adapted and one non-adapted. Since the game was specifically developed for this study, it was novel to all participants. While the game included similar game mechanics as Beat Saber and other VR rhythm dance games, it introduced new game mechanics to dynamically adapt difficulty during gameplay. Consequently, even participants with experience in rhythm dance games had to familiarize themselves with the new gameplay. It is likely that participants may have not fully understood all game mechanics when playing the first or even both games, and thus, results may differ if the game was played over a longer period of time.

Additionally, the game incorporated two bonus blocks, each resulting in different effects and varying impact. The star bonus was more beneficial than the attack bonus but spawned further away. Consequently, some participants may have preferred the attack blocks more often as they were easier to reach. While the star bonus aided in scoring significantly higher, attacking the opponent made it challenging for their opponent to score. Ultimately, the star bonus proved more efficient in overcoming large score differences. Therefore, the score difference might have been even smaller if participants had a complete understanding of the game mechanics, allowing more strategic gameplay.

7.3.3 Novelty Effect

The survey results revealed that, with the exception of two participants, most were either new to rhythm dance games or played them rarely. Additionally, this specific exergame introduced novel game mechanics, such as attacking opponents and dynamic difficulty adaptation. Consequently, a novelty effect was likely present, enhancing the game's interest for participants. Therefore, it is plausible that the results, particularly the scores from the motivation questionnaires, may have been influenced by this novelty effect.

7.3.4 Issues During Gameplay

Several issues emerged during gameplay that might have impacted the study's results. One participant, who wore glasses, opted not to wear them due to discomfort, resulting in an inability to read some in-game texts and, consequently, a diminished gaming experience. While other participants with glasses chose to wear them, they might have experienced some discomfort as well, potentially affecting their game experience.

Another participant misunderstood the explanation of the game mechanics, specifically regarding hitting blocks in the right direction, leading to numerous missed blocks initially. This participant only played the game correctly in the second round. Additionally, three participants either pressed a button or stepped outside the designated boundary during gameplay, requiring assistance to continue. As a result, they missed more blocks, leading to lower scores than they might have achieved. All these issues could have introduced biases into the participants' responses, given their encountered problems during the gameplay.

7.3.5 Fairness Questionnaire

As there is no evaluated questionnaire about fairness in multiplayer games, statements were created to ask about fairness. The evaluation of the fairness questionnaire was not part of this thesis. Therefore, the questionnaire might not be valid and reliable.

7.4 Further Research

As presented in the section about limitations, certain biases could have influenced the results. Furthermore, it is not clear yet if players could be motivated to play exergames with dynamic difficulty adaptation long term, and if the derived requirements can be applied to other types of exergames. Additionally, the interviews have shown that the game can still further improve, especially in user experience, but also in the method of adapting difficulty. The following sections discuss the aforementioned aspects in light of conducting further research.

7.4.1 Experiment Context

In this study, motivation is emphasized while playing the game. Ultimately, the goal of dynamic difficulty adaptation is to encourage people to engage in exergames with friends over an extended period, as frustration can be mitigated.

As the user experiments only included one session with two games, it was too short to examine if players prefer to play with dynamic difficulty adaptation over a long period. Therefore, further research could delve into studies covering multiple weeks in a more realistic environment, where players play at home remotely with friends. It could be examined how often players choose to play exergames with dynamic difficulty adaptation compared to without.

Another aspect to examine is in what context dynamic difficulty adaptation is most useful. The interview results suggest that dynamic difficulty adaptation may be more beneficial for beginners in exergaming or casual gamers. Over a long-term study, it could be investigated whether beginner players are likely to choose exergames with dynamic difficulty adaptation initially and if they will opt to play without it after gaining more experience.

7.4.2 Applying Findings to Other Exergames

This research focused on dance rhythm exergames, and focus group participants who collaboratively developed solutions for incorporating dynamic difficulty adaptation concentrated on applying it to this specific type of exergame. Thus, the requirements derived from the research results may not be applicable to other types of exergames. While the requirements have been formulated purposefully generic to be applied to many types of exergames, other types may require different approaches to dynamic difficulty adaptation.

For example, in boxing exergames, players are directly competing against each other, meaning that they can see their opponents' movements directly, while their movements are observed by their opponent as well. Therefore, players may feel more exposed when they receive assistance if their performance is weaker. In this research's game, players might have felt less observed, and as a result, they did not have a problem with the visually explicit method to adapt difficulty.

Furthermore, this research focuses on exergames with two players playing against each other. In exergames with multiple players, the dynamics between players may change, and thus, dynamic difficulty adaptation may have different effects or need to be applied differently to achieve positive effects.

7.4.3 Further Game Improvements

Participants stated in interviews that they did not always understand how the scoring worked in the game. For example, multiple factors could lead to an increase in score, but it was difficult for participants to grasp. Furthermore, certain attacks were too extreme and became frustrating, especially in the adapted version of the game, due to many consecutive attacks. The game should give players the ability to overcome attacks.

Participants could not pay attention to all elements of the game that were visually present in the game due to the high pace of the game. Exergames most likely have a high pace, and therefore, it has to be carefully decided what information is shown during the game and in what form. Thus, only information that is crucial for game play should be shown. The loading bars to indicate how much a player moves also included loading factors when the loading bar was loading faster. The participants did not pay attention to this and did not fully understand what it meant. It was also not crucial for the game and therefore should have been omitted. This information was only included to create full transparency of the dynamic difficulty adaptation. But it became apparent during the user experiments that players will not take into account information that they do not influence themselves. Furthermore, presenting information as text is difficult to process during the game and therefore should be avoided.

Some ideas for including further information about true skills after the game could be added to make players aware of their progress and create another way to motivate people to play exergames over a long period. With the aforementioned improvements to the game, another study could elaborate on the effects of dynamic difficulty adaptation, as these issues may have influenced the results. Ideally, an improved game would be used in a long-term study, as mentioned in a previous section, to investigate the effects of long-term motivation.

8 Conclusion

Exergames have the potential to motivate people to move regularly in the space of their own home, requiring less time commitment. Playing with friends can further motivate players to engage in long-term physical activity and acts as another way to connect with friends remotely. However, friends may have varying degrees of skill, as they may play exergames or engage in physical activities more or less frequently. Thus, it can be frustrating to play against friends who often win or boring to play against those who frequently lag behind. This thesis explores dynamically adapting difficulty according to each player's performance as a potential solution to this issue. To examine the effects of dynamic difficulty adaptation, the following research question was formulated:

How can dynamic difficulty adaptation be used to balance mismatching skills of multiple players?

This research question was answered by deriving requirements from user discussions as well as investigating the created design of dynamic difficulty adaptation in terms of competence, motivation and fairness. It was hypothesized that dynamically adapting an exergame's difficulty would have positive effects on the player's perceived competence, motivation, and fairness of the game. To develop an exergame with dynamic difficulty adaptation that would have positive effects, the thesis followed the Design Science framework. The first part of the framework involved developing a prototype based on the needs of the player. Specifically, the focus was on rhythm dance games to allow a specific solution to be designed. Through discussions in focus groups, requirements for dynamic difficulty adaptation were derived. These requirements were then used in subsequent focus groups to further elaborate on a specific design. The resulting design concept was implemented as a prototype.

The second part of the framework focused on the empirical study. Using the prototype, user experiments were conducted to investigate whether the hypotheses were true and if the earlier requirements could lead to a successful solution that balances the exergame and is accepted among players.

The results show that dynamically adapting difficulty based on players' performances and score differences led to a significant higher perception of fairness. Participants pointed out that the ability to catch up at almost any point in time during the game encourages them to put effort into playing. Skilled players also feel more challenged by their opponents due to dynamic difficulty adaptation. Even though some participants stated that the adapted game motivated them, there was no significant difference regarding motivation between the adapted and non-adapted game. The same applies to perceived competence during the games. Some participants also felt frustrated due to many attacks that were too extreme. While the adaptation gave weaker players ways to catch up, it may have also increased the difficulty for better players. However, increasing difficulty in the game needs to be carefully designed, as it should still adhere to Flow theory in order to not overwhelm players but challenge them.

Previous research on balancing exergames has found that explicit balancing methods can lead to a decline in self-efficacy in players. This specifically happened to weaker players as they felt exposed when receiving assistance. However, this was not examined in the conducted user experiments. A reason for this could be that participants found the adapted game more difficult as the number of blocks spawned increased greatly. Furthermore, each player was concentrated on their own blocks and therefore did not observe their opponents nor felt themselves observed. Therefore, players may only feel exposed with explicit difficulty balancing when they are observed by their opponents.

Some participants preferred the non-adapted game as they would like to compare true skills with other players. This aligns with findings of the focus groups: very competitive players are seeking to measure and compare skills and therefore do not see value in balancing exergames by adjusting difficulties for individual players.

The comparison of the derived requirements of the focus groups with the interview results showed that most requirements were successfully contributing to a balanced game. These requirements can be taken into account by game designers when implementing an approach to dynamically adapt difficulty in exergames.

The user study also had some limitations. For example, the participant sample was too small to get reliable results for the combination of various variables. The experiment context was not realistic as players were in the same room and only played two games in one session. As this game was created for the user study, participants were new to the game and did not have much time to learn it. Furthermore, most participants were new to rhythm dance games, which may have led to a novelty effect that influenced results. Some participants also experienced issues during the game, which may have also influenced their game experience and their answers.

This research showed that there is potential for dynamic difficulty adaptation. As assumed, it creates a fair game, where both players receive assistance when needed. This can motivate players throughout the game as they feel that they can still win by putting more effort into playing. Future research can build on these results and focus on the effect on long-term motivation, as well as explore different use cases. For example, it could investigate whether casual players or beginner players are more likely to accept balancing exergames. Furthermore, this research gives insight on how to design dynamic difficulty adaptation in an exergame according to players' needs and expectations. A future study should take place in a realistic setting, meaning that participants play exergames remotely in their homes over the course of a long period.

Bibliography

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. 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* 135 (2020), 102370.

[2] Clark C Abt. 1987. *Serious games.* University press of America.

[3] David Altimira, Mark Billinghurst, and Florian Mueller. 2013. Understanding handicapping for balancing exertion games. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems.* 1125–1130.

[4] David Altimira, Florian'Floyd' Mueller, Gun Lee, Jenny Clarke, and Mark Billinghurst. 2014. Towards understanding balancing in exertion games. In *Proceedings of the 11th conference on advances in computer entertainment technology.* 1–8.

[5] Julian Alvarez, Damien Djaouti, et al. 2011. An introduction to Serious game Definitions and concepts. *Serious games & simulation for risks management* 11, 1 (2011), 11–15.

[6] Cay Anderson-Hanley, Amanda L Snyder, Joseph P Nimon, and Paul J Arciero. 2011. Social facilitation in virtual reality-enhanced exercise: competitiveness moderates exercise effort of older adults. *Clinical interventions in aging* (2011), 275–280.

[7] James Annesi. 2003. Effects of a cognitive behavioral treatment package on exercise attendance and drop out in fitness centers. *European Journal of Sport Science* 3, 2 (2003), 1–16.

[8] James J Annesi. 1998. Effects of computer feedback on adherence to exercise. *Perceptual and motor skills* 87, 2 (1998), 723–730.

[9] Fabio Antonialli, Luiz Marcelo Antonialli, and Renan Antonialli. 2017. Uses and abuses of the Likert scale: bibliometric study in the proceedings of Enanpad from 2010 to 2015. *Revista Reuna* 22, 4 (2017), 1–19.

[10] Oluleke Bamodu and Xu Ming Ye. 2013. Virtual reality and virtual reality system components. In *Advanced materials research*, Vol. 765. Trans Tech Publ, 1169–1172.

[11] Rosa María Baños, Cristina Botella, Azucena Garcia-Palacios, Helena Villa, Concepción Perpiñá, and Mariano Alcaniz. 2000. Presence and reality judgment in virtual environments: a unitary construct? *CyberPsychology & Behavior* 3, 3 (2000), 327–335.

[12] Brett Barron and David Chorney. 2016. Perceptions of Multiplayer Exergames Within PE: A Mixed Method Study. *Think Like a Mountain Comprehensive School Health in Six Priority Areas* (2016), 41.

[13] Roy F Baumeister and Mark R Leary. 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin* 117, 3 (1995), 497.

[14] Aslihan Tece Bayrak, Rahul Kumar, Jak Tan, DeVon AhMu, Jordan Hohepa, Lindsay A Shaw, Christof Lutteroth, and Burkhard C Wünsche. 2017. Balancing different fitness levels in competitive exergames based on heart rate and performance. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*. 210–217.

[15] Nadia Bianchi-Berthouze, Whan Woong Kim, and Darshak Patel. 2007. Does body movement engage you more in digital game play? and why?. In *Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007 Proceedings 2*. Springer, 102–113.

[16] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI research: Going behind the scenes*. Morgan & Claypool Publishers.

[17] Susanne Bødker, Christian Dindler, Ole S Iversen, and Rachel C Smith. 2022. What Is Participatory Design? In *Participatory Design*. Springer, 5–13.

[18] Doug A. Bowman and Ryan P. McMahan. 2007. Virtual Reality: How Much Immersion Is Enough? *Computer* 40, 7 (2007), 36–43. https://doi.org/10.1109/MC.2007.257

[19] Johannes Breuer and Gary Bente. 2010. Why so serious? On the relation of serious games and learning. *Journal for Computer Game Culture* 4 (2010), 7–24.

[20] Ed Burke. 1998. *Precision heart rate training*. Human Kinetics.

[21] Barbara Bushman, American College of Sports Medicine, et al. 2017. *ACSM's Complete Guide to Fitness & Health, 2E*. Human Kinetics.

[22] Juliet M Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1 (1990), 3–21.

[23] Kevin Corti. 2006. Games-based Learning; a serious business application. *Informe de PixelLearning* 34, 6 (2006), 1–20.

[24] Cora L Craig, Alison L Marshall, Michael Sjöström, Adrian E Bauman, Michael L Booth, Barbara E Ainsworth, Michael Pratt, ULF Ekelund, Agneta Yngve, James F Sallis, et al. 2003. International physical activity questionnaire: 12-country reliability and validity. *Medicine & science in sports & exercise* 35, 8 (2003), 1381–1395.

[25] Mihaly Csikszentmihalyi. 1990. *Flow: The Psychology of Optimal Experience*. Harper Perennial.

[26] Edward L Deci. 1992. The relation of interest to the motivation of behavior: A self-determination theory perspective. *The role of interest in learning and development* 44 (1992).

[27] Ansgar E Depping, Regan L Mandryk, Chengzhao Li, Carl Gutwin, and Rodrigo Vicencio-Moreira. 2016. How disclosing skill assistance affects play experience in a multiplayer first-person shooter game. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 3462–3472.

[28] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness: defining" gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. 9–15.

[29] Rod K Dishman. 1988. *Exercise adherence: Its impact on public health*. Human Kinetics.

[30] Damien Djaouti, Julian Alvarez, Jean-Pierre Jessel, and Olivier Rampnoux. 2011. Origins of serious games. *Serious games and edutainment applications* (2011), 25–43.

[31] Nuša Farič, Lee Smith, Adrian Hon, Henry WW Potts, Katie Newby, Andrew Steptoe, Abi Fisher, et al. 2021. A virtual reality exergame to engage adolescents in physical activity: Mixed methods study describing the formative intervention development process. *Journal of medical Internet research* 23, 2 (2021), e18161.

[32] Canadian Fitness and Lifestyle Research Institute. 2005. *2005 Physical Activity Monitor*.

[33] Cuisle Forde. 2018. Scoring the international physical activity questionnaire (IPAQ). *University of Dublin* 3 (2018).

[34] Kennon Francis. 1996. Physical activity in the prevention of cardiovascular disease. *Physical therapy* 76, 5 (1996), 456–468.

[35] Kathrin Maria Gerling, Matthew Miller, Regan L Mandryk, Max Valentin Birk, and Jan David Smeddinck. 2014. Effects of balancing for physical abilities on player performance, experience and self-esteem in exergames. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2201–2210.

[36] Frederick X Gibbons and Bram P Buunk. 1999. Individual differences in social comparison: development of a scale of social comparison orientation. *Journal of personality and social psychology* 76, 1 (1999), 129.

[37] Stefan Göbel, Sandro Hardy, Viktor Wendel, Florian Mehm, and Ralf Steinmetz. 2010. Serious games for health: personalized exergames. In *Proceedings of the 18th ACM international conference on Multimedia*. 1663–1666.

[38] Johanna Höysniemi. 2006. *Design and evaluation of physically interactive games*. Tampere University Press.

[39] Susan Hwang, Adrian L Jessup Schneider, Daniel Clarke, Alexander Macintosh, Lauren Switzer, Darcy Fehlings, and TC Nicholas Graham. 2017. How game balancing affects play: Player adaptation in an exergame for children with cerebral palsy. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 699–710.

[40] Susan Jamieson. 2004. Likert scales: How to (ab) use them? *Medical education* 38, 12 (2004), 1217–1218.

[41] Mads Møller Jensen and Kaj Grønbæk. 2016. Design strategies for balancing exertion games: A study of three approaches. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 936–946.

[42] Maximus D Kaos, Ryan E Rhodes, Perttu Hämäläinen, and TC Nicholas Graham. 2019. Social play in an exergame: how the need to belong predicts adherence. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.

[43] Tanja Kojić, Jan-Niklas Voigt-Antons, Steven Schmidt, Lukas Tetzlaff, Bruno Kortowski, Uliana Sirotina, and Sebastian Möller. 2018. Influence of virtual environments and conversations on user engagement during multiplayer exergames. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 1–3.

[44] Sujit Kumar Basak, Marguerite Wotto, and Paul Belanger. 2018. E-learning, M-learning and D-learning: Conceptual definition and comparative analysis. *E-learning and Digital Media* 15, 4 (2018), 191–216.

[45] Fedwa Laamarti, Mohamad Eid, and Abdulmotaleb El Saddik. 2014. An overview of serious games. *International Journal of Computer Games Technology* 2014 (2014), 11–11.

[46] Elizabeth J Lyons. 2015. Cultivating engagement and enjoyment in exergames using feedback, challenge, and rewards. *Games for health journal* 4, 1 (2015), 12–18.

[47] Alexander MacIntosh, Lauren Switzer, Susan Hwang, Adrian L Jessup Schneider, Daniel Clarke, TC Nicholas Graham, and Darcy L Fehlings. 2017. Ability-based balancing using the gross motor function measure in exergaming for youth with cerebral palsy. *Games for Health Journal* 6, 6 (2017), 379–385.

[48] Andrew Macvean and Judy Robertson. 2013. Understanding exergame users' physical activity, motivation and behavior over time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1251–1260.

[49] Arwen M Marker and Amanda E Staiano. 2015. Better together: outcomes of cooperation versus competition in social exergaming. *Games for health journal* 4, 1 (2015), 25–30.

[50] Anna Lisa Martin-Niedecken. 2018. Designing for bodily interplay: engaging with the adaptive social exertion game" plunder planet". In *Proceedings of the 17th ACM conference on interaction design and children*. 19–30.

[51] Amir Matallaoui, Jonna Koivisto, Juho Hamari, and Ruediger Zarnekow. 2017. How effective is "exergamification"? A systematic review on the effectiveness of gamification features in exergames. (2017).

[52] Edward McAuley, Terry Duncan, and Vance V Tammen. 1989. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport* 60, 1 (1989), 48–58.

[53] David R Michael and Sandra L Chen. 2005. *Serious games: Games that educate, train, and inform.* Muska & Lipman/Premier-Trade.

[54] Mahmoud Mohanna. 2015. *Using knowledge engineering for modeling mobile learning systems.* Ph. D. Dissertation. Université Laval.

[55] Florian Mueller, Frank Vetere, Martin Gibbs, Darren Edge, Stefan Agamanolis, Jennifer Sheridan, and Jeffrey Heer. 2012. Balancing exertion experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 1853–1862.

[56] Jeanne Nakamura and Mihaly Csikszentmihalyi. 2009. Flow theory and research. *Handbook of positive psychology* 195 (2009), 206.

[57] David L Neumann, Robyn L Moffitt, Patrick R Thomas, Kylie Loveday, David P Watling, Chantal L Lombard, Simona Antonova, and Michael A Tremeer. 2018. A systematic review of the application of interactive virtual reality to sport. *Virtual Reality* 22 (2018), 183–198.

[58] Nintendo. [n. d.]. *Creating smiles for generations - Our history.* https://www.nintendo.com/about/

[59] Nurshamine Nor, Mohd Sunar, and Azyan Kapi. 2019. A review of gamification in Virtual Reality (VR) sport. *EAI Endorsed Transactions on Creative Technologies* 6, 21 (2019).

[60] Nurshamine Nazira Nor, Mohd Shahrizal Sunar, and Azyan Yusra Kapi. 2020. User experience of gamified virtual reality (VR) in sport: A review. *Science and Technologies for Smart Cities: 5th EAI International Summit, SmartCity360, Braga, Portugal, December 4-6, 2019, Proceedings* (2020), 440–449.

[61] Mateus Nunes, Luciana Nedel, and Valter Roesler. 2014. Motivating people to perform better in exergames: Competition in virtual environments. In *Proceedings of the 29th annual ACM symposium on applied computing*. 970–975.

[62] Jennifer A O'dea. 2003. Why do kids eat healthful food? Perceived benefits of and barriers to healthful eating and physical activity among children and adolescents. *Journal of the American Dietetic Association* 103, 4 (2003), 497–501.

[63] World Health Organization. 2000. Obesity: preventing and managing the global epidemic. (2000).

[64] Wei Peng and Julia Crouse. 2013. Playing in parallel: The effects of multiplayer modes in active video game on motivation and physical exertion. *Cyberpsychology, behavior, and social networking* 16, 6 (2013), 423–427.

[65] Mikki H Phan, Jo R Jardina, Sloane Hoyle, and Barbara S Chaparro. 2012. Examining the role of gender in video game usage, preference, and behavior. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 56. Sage Publications Sage CA: Los Angeles, CA, 1496–1500.

[66] Marc Prensky. 2001. Digital natives, digital immigrants part 2: Do they really think differently? *On the horizon* 9, 6 (2001), 1–6.

[67] Alejandro Quintas-Hijós, Carlos Peñarrubia-Lozano, and Juan Carlos Bustamante. 2020. Analysis of the applicability and utility of a gamified didactics with exergames at primary schools: Qualitative findings from a natural experiment. *PloS one* 15, 4 (2020), e0231269.

[68] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[69] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30 (2006), 344–360.

[70] Ben Sawyer. 2002. Serious games: Improving public policy through game-based learning and simulation. (01 2002).

[71] Simone M Schneider and Jürgen Schupp. 2014. Individual differences in social comparison and its consequences for life satisfaction: introducing a short scale of the

Iowa–Netherlands Comparison Orientation Measure. *Social Indicators Research* 115 (2014), 767–789.

[72] Sebastián López Serrano, Alberto Ruiz-Ariza, Manuel De La Torre-Cruz, and Emilio José Martínez López. 2021. Improving cognition in school children and adolescents through exergames. A systematic review and practical guide. *South African Journal of Education* 41, 1 (2021).

[73] Lindsay A Shaw, Jude Buckley, Paul M Corballis, Christof Lutteroth, and Burkhard C Wünsche. 2016. Competition and cooperation with virtual players in an exergame. *PeerJ Computer Science* 2 (2016), e92.

[74] Jeff Sinclair, Philip Hingston, and Martin Masek. 2007. Considerations for the design of exergames. In *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*. 289–295.

[75] Amanda L Snyder. 2011. Virtual and live social facilitation while exergaming: Competitiveness moderates. (2011).

[76] Hayeon Song, Jihyun Kim, Kelly E Tenzek, and Kwan Min Lee. 2013. The effects of competition and competitiveness upon intrinsic motivation in exergames. *Computers in Human Behavior* 29, 4 (2013), 1702–1708.

[77] Tadeusz Stach, TC Nicholas Graham, Jeffrey Yim, and Ryan E Rhodes. 2009. Heart rate control of exercise video games. In *Proceedings of Graphics interface 2009*. 125–132.

[78] Amanda E Staiano, Anisha A Abraham, and Sandra L Calvert. 2013. Adolescent exergame play for weight loss and psychosocial improvement: a controlled physical activity intervention. *Obesity* 21, 3 (2013), 598–601.

[79] Amanda E Staiano and Sandra L Calvert. 2011. Exergames for physical education courses: Physical, social, and cognitive benefits. *Child development perspectives* 5, 2 (2011), 93–98.

[80] Jonathan Steuer, Frank Biocca, Mark R Levy, et al. 1995. Defining virtual reality: Dimensions determining telepresence. *Communication in the age of virtual reality* 33 (1995), 37–39.

[81] Tarja Susi, Mikael Johannesson, and Per Backlund. 2007. Serious games: An overview. (2007).

[82] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)* 3, 3 (2005), 3–3.

[83] Christian Swertz. 2009. Serious games taken seriously. *Eludamos: journal for computer game culture* 3, 1 (2009), 7–8.

[84] Áron Tóth and Emma Lógó. 2018. The effect of gamification in sport applications. In *2018 9th IEEE international conference on cognitive infocommunications (coginfocom)*. IEEE, 000069–000074.

[85] Elena Tuveri, Luca Macis, Fabio Sorrentino, Lucio Davide Spano, and Riccardo Scateni. 2016. Fitmersive games: Fitness gamification through immersive VR. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. 212–215.

[86] Lotte Vermeulen, Jan Van Looy, Frederik De Grove, and Cédric Courtois. 2011. You are what you play?: A quantitative study into game design preferences across gender and their interaction with gaming habits. In *DiGRA 2011: Think, design, play*. Digital Games Research Association (DiGRA).

[87] Rodrigo Vicencio-Moreira, Regan L Mandryk, and Carl Gutwin. 2015. Now you can compete with anyone: Balancing players of different skill levels in a first-person shooter game. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2255–2264.

[88] G. Vikas and K. Vineet. 2022. *Serious Games Market by Gaming Platform (Smartphone, Console, PC, and Others), Application (Simulation and Training, Research and Planning, Advertising and Marketing, Human Resources, and Others), Industry Vertical (Education, Healthcare, Aerospace and Defense, Government, Retail, Media and Entertainment, and Others), and Region: Global Opportunity Analysis and Industry Forecast, 2021-2030.* https://www.alliedmarketresearch.com/serious-games-market

[89] John CK Wang, Woon Chia Liu, Marc R Lochbaum, and Sarah J Stevenson. 2009. Sport ability beliefs, 2 x 2 achievement goals, and intrinsic motivation: The moderating role of perceived competence in sport and exercise. *Research Quarterly for Exercise and Sport* 80, 2 (2009), 303–312.

[90] Roel J Wieringa. 2014. *Design science methodology for information systems and software engineering*. Springer.

Appendix

A Protocols

A.1 Preliminary Focus Groups Protocol

Phase	Facilitator task	Aim	Time
Opening the focus group	Greet participants Thank them for participating Explain in short the aim of the session and research: "We want to explore how multiplayer exergames in VR between players with different skills can be made fair and motivating. Multiple players play exergames together in VR but have different abilities and skills. This can lead to frustration due to one player dominating the games. However, we want to motivate players to keep playing exergames to encourage physical activities and avoid frustration." Handover consent form	Establish a comfortable environment for all participants to maximise the creativity and freedom of expression. All participants get time to read through the documents and sign the consent form.	5 min
Clarification before start	Ask again about consent for recordings Ask for any questions	The participants use this phase to ask questions about the research.	2-4 min
1. Question	Remind of scenario: "Multiple players with different skills play together but one or more players are clearly dominating and almost always winning. How can we create a better game experience for both players that does not create frustration?".	Gather participants' ideas how they would solve the problem of mismatching skills between players and resulting frustrations.	5-7 min

2. Question	Explain dynamic difficulty adaptation and why it could be a potential solution. "What do you think about adapting difficulty dynamically for players to create an equal chance of winning between players?".	Discussion about dynamic difficulty adaptation.	5-7 min
3. Question	"What types of difficulty adaptation would you like to see in exergames? What would you not like to see?".	Gather participants' ideas of how to adapt difficulty.	5-7 min
4. Question	Explain the difference between static and dynamic difficulty adaptation. "Should static or dynamic difficulty adaptation be used?".	Discussions about different types of difficulty adaptation.	5-7 min
5. Question	"Should be difficulty rather increased or decreased or both?".	Discussions about the direction of difficulty adapation (increase vs. decrease).	5-7 min
6. Question	"Should both players or only one player experience the difficulty adaptation?".	Discussions about who should receive assistance.	5-7 min
Close focus group	Explain how they helped in the process of this research and how the next steps will of the research will look like. Thank them for participation Let them know that they can leave their email address in case they would like to receive the results of this research when it is finished.	End the focus group.	2 min

Table 10: Preliminary focus group protocol

A.2 Design Focus Group Protocol

A.2.1 1. Design Focus Group Protocol

Phase	Facilitator task	Aim	Time
Opening the focus group	Greet participants Thank them for participating Explain in short the aim of the session and research: "We want to explore how multiplayer exergames in VR between players with different skills can be made fair and motivating. Multiple players play exergames together in VR but have different abilities and skills. This can lead to frustration due to one player dominating the games. However, we want to motivate players to keep playing exergames to encourage physical activities and avoid frustration." Explain that their goal today is to create a simple paper prototype showing a way to solve this particular problem. Handover consent form	Establish a comfortable environment for all participants to maximise the creativity and freedom of expression. All participants get time to read through the documents and sign the consent form.	5 min

Task explanation	Present what insights have been gathered beforehand about this topic. These insights should be taken into account when designing a solution. Hand them over a summary of these insights, which they can take a look on while designing. Ask again about consent for recordings Ask for any questions	The participants get further insights into the task and all necessary information are presented that are needed to solve the task. Participants use this phase to ask questions about the research.	5 min
1. Iteration	Explain an example VR exergame that is popular: Beat Saber. Show a video of the gameplay and offer the possibility to wear VR glasses to see the game. "How would you design a fair and dynamic difficulty adaptation for the game Beat Saber? Please be aware here that the focus is to make the solution as fair as possible for all players".	This research will focus on exergames that have the same or similar game mechanic as the mentioned game. The participants should now think of ways, how they would design difficulty adaptation for this particular game while keeping in mind that the gameplay should be as fair as possible.	15 min
1. Presentation & discussion	Ask participants to pitch their ideas quickly to the rest of the group. After each short presentation, ask the participants to point out strengths and weaknesses of the solution.	After the first iteration, participants should reflect on all solutions.	6-9 min

2. Iteration	Explain the problem that players of exergames might move less to gain points more efficiently, which is contradicting when exergames should encourage players to move. This problem exists with Beat Saber as well. "How would you encourage players of exergames to move more in this game?". Encourage participants to use the insights of the previous discussions to improve their prototype. It is allowed to use parts of other solutions that they liked.	Participants can now improve their prototype and at the same time incorporate a way to encourage players to move more.	15 min
2. Presentation & discussion	Ask participants to pitch their final prototype shortly to the rest of the group. After all participants presented, ask them to discuss finally how they would provide a personalised solution. This means, how can a solution provided that works for all players of different personalities and preferences. For example, can players choose certain aspects of the solution? How would the solutions be combined?	Final discussions of the presented solutions to gather their opinions of which solution they liked best and how to combine the solutions to provide a personalised solution.	10 min

Close focus group	Explain how they helped in the process of this research and how the next steps will of the research will look like. Thank them for participation Let them know that they can leave their email address in case they would like to receive the results of this research when it is finished.	End the focus group.	2 min

Table 11: Design focus group protocol

A.2.2 2. Design Focus Group Protocol

Phase	**Facilitator task**	**Aim**	**Time**

Opening the focus group	Greet participants Thank them for participating Explain in short the aim of the session and research: "We want to explore how multiplayer exergames in VR between players with different skills can be made fair and motivating. Multiple players play exergames together in VR but have different abilities and skills. This can lead to frustration due to one player dominating the games. However, we want to motivate players to keep playing exergames to encourage physical activities and avoid frustration. For this, we propose dynamic difficulty adaptation as a solution. Difficulty is therefore always adapted to create a close game between the players, which means weaker players get chances to catch up and stronger players get more challenged during the game". Explain that they will be presented with four designs of difficulty adaptation in a rhythm exergame, which they will discuss. The rhythm exergame is similar to Beat Saber. Show a video of Beat Saber and explain the game mechanics. Explain that they should discuss the ideas keeping in mind that the game should be fair and motivating for players with mismatching skills. Handover consent form Ask again about consent for recordings Ask for any questions	Establish a comfortable environment for all participants to maximise the creativity and freedom of expression. All participants get time to read through the documents and sign the consent form.	8 min

1. Design	Idea: Create power-ups in the game that players can activate, for example by hitting a specific object. These power-ups are advantages that help players to gain more scores for a short period of time. For example: clearing (hitting) all objects at once, switching sabers with pistols to hit targets from far away, higher scores for all targets etc. Both players receive power-ups but the weaker player receives more. What power-ups would you create?	Participants discuss the first idea.	https://www.overl 5-10 min
2. Design	Idea: Players can mess with the game of the opponent. By hitting a specific target they can increase the difficulty of their opponent for a short time. For example, the targets are more spread out in distance, which makes it harder to reach or targets are significantly smaller or less score can be made. Weaker players get more possibilities to mess with the opponent's game than stronger players. How would you increase the difficulty of the opponent?	Participants discuss the second idea.	5-10 min
3. Design	Idea: The idea is to create a game similar to *Twister*. Players have to recreate poses which get displayed for bonus points. Weaker players get more possibilities to do so than stronger players.	Participants discuss the third idea.	5-10 min

4. Design	Idea: Automatically increase or decrease the size of objects in the game in order to make it more or less difficult for both players. For example, the targets could get bigger or smaller. The sabers could change. Here the player does not activate the adaptation. Instead it happens automatically. Do you see problems with this approach? Do you see advantages with this approach?	Participants discuss the fourth idea.	5-10 min
Last questions	Which idea did you like best and why? Would you prefer to combine ideas? Which ones would you combine?	Participants explain their favorite idea.	5-10 min
Close focus group	Explain how they helped in the process of this research and how the next steps will of the research will look like. Thank them for participation Let them know that they can leave their email address in case they would like to receive the results of this research when it is finished.	End the focus group.	2 min

Table 12: Second Design focus group protocol

A.3 User Study Protocol

Phase	**Researcher task**	**Aim**

Introduction	Greet participants Thank them for participating Explain in short the aim of this research: ”We want to explore how multiplayer exergames in VR between players with different skills can be made fair and motivating. Multiple players play exergames together in VR but have different abilities and skills. This can lead to frustration due to one player dominating the games. However, we want to motivate players to keep playing exergames to encourage physical activities and avoid frustration.” Handover consent form as web survey	Establish a comfortable environment for both participants. Give them time to read the consent form in their own pace.
Clarification before start	Ask again about consent for recordings Explain basic game rules. Attach heart rate monitors to participants and not down initial heart rate before playing. Ask for any questions	The participants use this phase to ask questions about the research.
First condition	Participants play a 3-4 min game together in one condition. Ask participants to fill out a questionnaire each.	Participants play one condition.
Second condition	Participants play a 3-4 min game together in the other condition. Ask participants to fill out a questionnaire each.	Participants play the other condition.
Short interview	Both participants are asked to answer 4 questions in a short interview. Audio is recorded.	Conducting short interview

Close user test	Explain how they helped in the process of this research and how the next steps will of the research will look like. Thank them for participation Let them know that they can leave their email address in case they would like to receive the results of this research when it is finished.	End the user test.

Table 13: User test protocol

B Focus Groups Consent Form

Universiteit Utrecht

Consent form for participation in the research project

Difficulty Adaptation in VR Multiplayer Exergames

Please complete the form below by ticking the relevant boxes and signing on the line below. A copy of the completed form will be given to you for your own record.

☐ I confirm that I am 18 years of age or over.

☐ I confirm that the research project ***"Difficulty Adaptation in VR Multiplayer Exergames"*** has been explained to me. I have had the opportunity to ask questions about the project and have had these answered satisfactorily. I had enough time to consider whether to participate.

☐ I consent to the material I contribute being used to generate insights for the research project "**Difficulty Adaptation in VR Multiplayer Exergames*".***

☐ I consent to audio recordings being used in this focus group. I understand that I can request to stop recordings at any time.

☐ I understand that if I give permission, the audio recordings will be held confidentially so that only Tamara Mantz have access to the recordings. The recordings will be held in a password protected storage for up to 3 months after which period they will be transcribed in an anonymous form and the original securely destroyed. In accordance with the General Data Protection Regulation (GDPR) I can have access to my recordings and can request them to be deleted at any time during this period.

☐ I understand that in addition to the recordings, other personal data will be collected from me and that this data will be held confidentially so that only Tamara Mantz has access to this data and are able to trace it back to me personally. The data will be held in a password protected storage for up to 3 months after which period it will be fully anonymized and securely destroyed. In accordance with the General Data Protection Regulation (GDPR) I can have access to my personal data and can request it to be deleted at any time during this period.

☐ I understand that my participation in this research is voluntary and that I may withdraw from the focus group at any time without providing a reason, and that if I withdraw any personal data already collected from me will be erased.

☐ I understand that my participation is not a requirement for my course or work, and that participating or not will not impact me.

☐ I consent to allow the fully anonymized data to be used in future publications and other scholarly means of disseminating the findings from the research project.

☐ I understand that the data acquired will be securely stored by researchers, but that appropriately anonymized data may in future be made available to others for research purposes. I understand that the University may publish appropriately anonymized data in appropriate data repositories for verification purposes and to make it accessible to researchers and other research users.

☐ I agree to take part in the above research project on ***"Difficulty Adaptation in VR Multiplayer Exergames"***.

Name of participant	Date	Signature

Tamara Mantz		
Name of researcher	Date	Signature

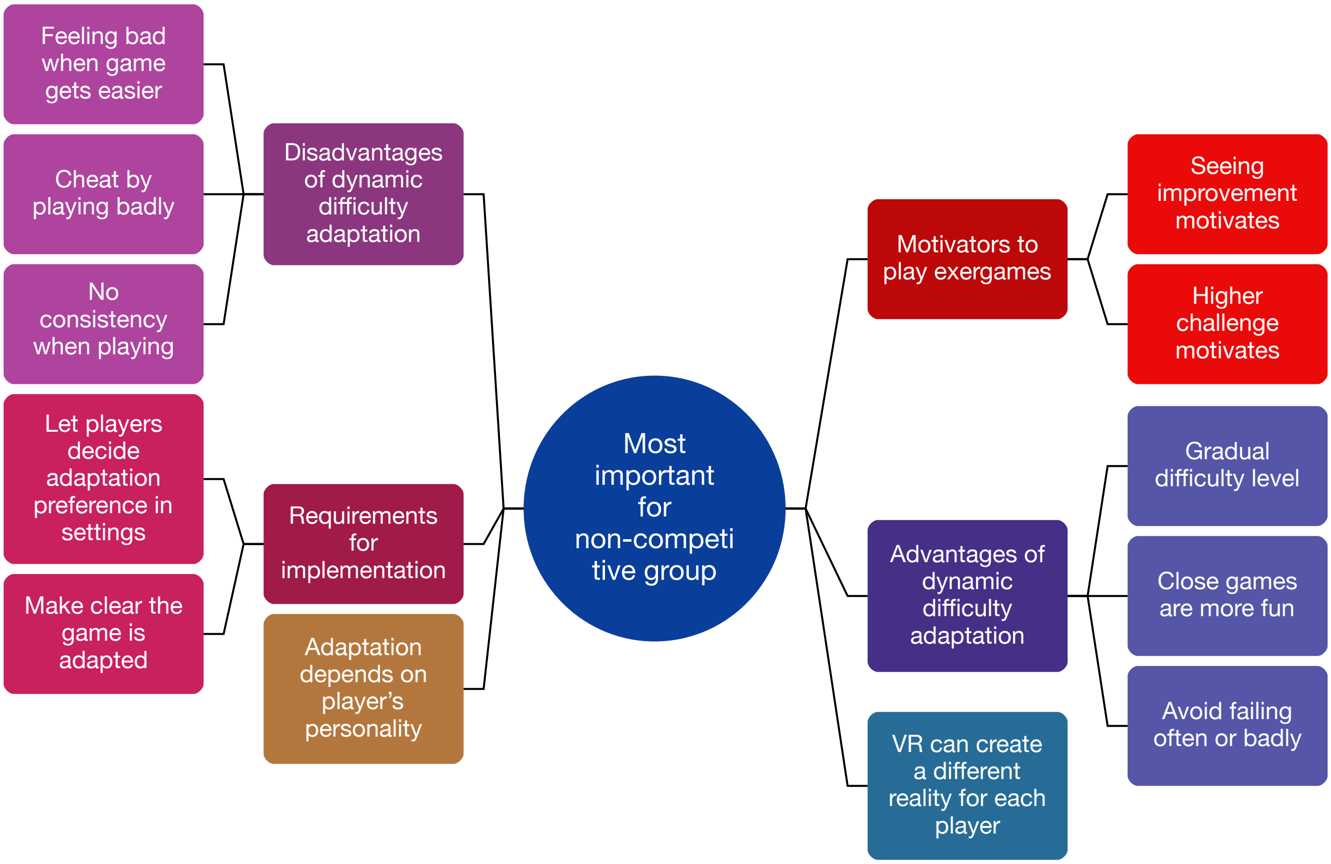
C Preliminary Focus Groups

C.1 Codes Sorted Into Categories

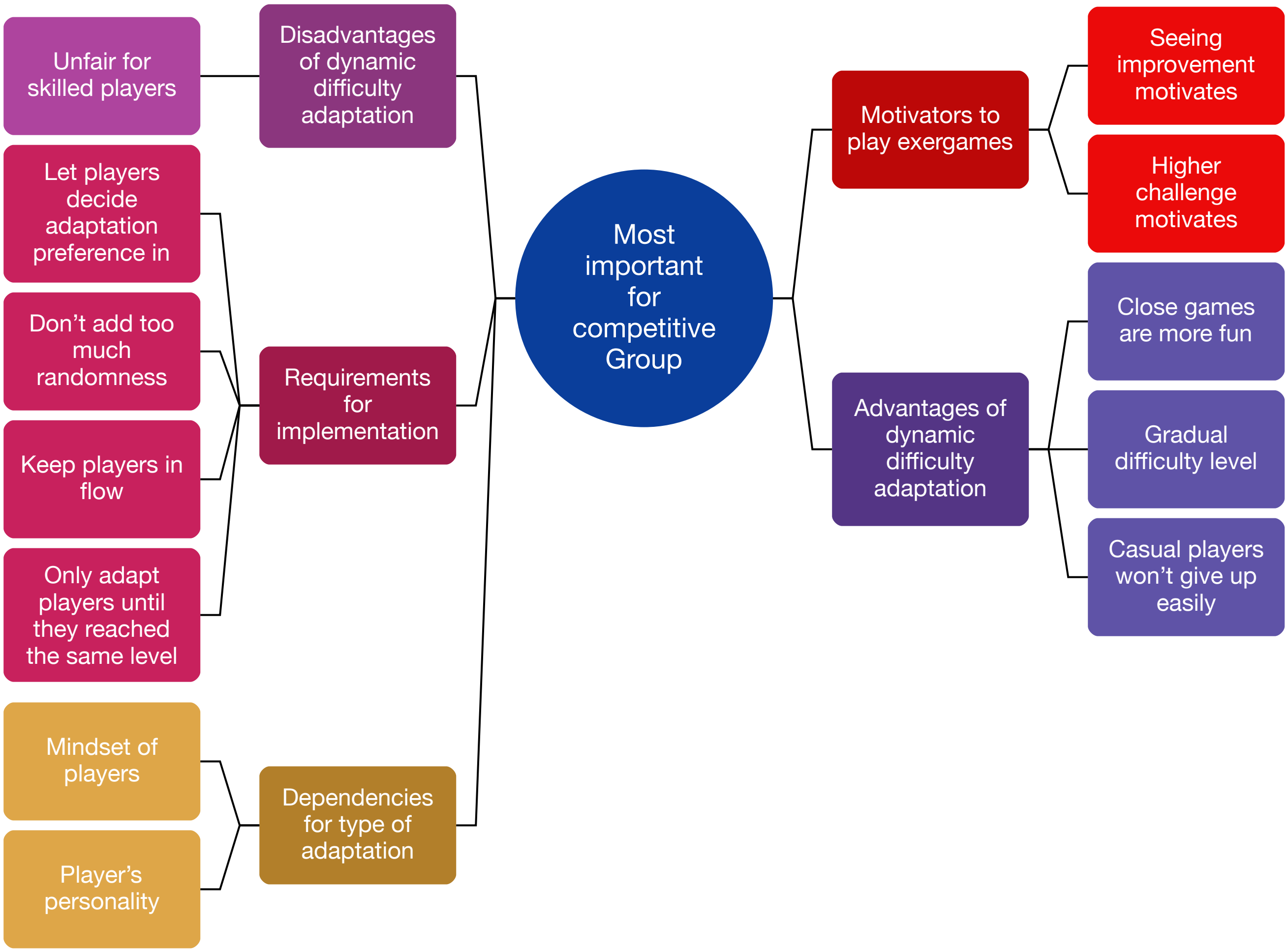
NC stands for non-competitive participant(s) and C for competitive participant(s).

Category	Code	# Statements	# Total in Category
Disadvantages of dynamic difficulty adaptation	Unfair for skilled players	12 (1NC, 11C)	30 (14NC, 16C)
	Feeling bad when game gets easier	5 (4NC, 1C)	
	Cheat by playing badly	4 (3NC, 1C)	
	No consistency when playing same game again	4 (3NC, 1C)	
	Makes players too equal	1 (1NC)	
	Removes challenges	1 (1NC)	
	Risk of wrong adaptation	1 (1NC)	
	Unfair due to less transparency	1 (1C)	
	Beating high score does not motivate necessarily	1 (1C)	
Motivators to play exergames	Seeing improvement motivates	15 (9NC, 6C)	25 (13NC, 12C)
	Higher challenge motivates	4 (2NC, 2C)	
	Reward players for playing again	2 (1NC, 1C)	
	Boost after winning against more skilled player	1 (1C)	
	Encourage with positive feedback	1 (1C)	
	Mentoring from skilled players	1 (1NC)	
	Seeing how skilled people play motivates	1 (1C)	
Requirements for implementation	Let players decide adaptation preference in settings	10 (4NC, 6C)	23 (10NC, 13C)
	Make clear the game is adapted	4 (4NC)	
	Don't add too much randomness	3 (3C)	
	Keep players in flow	3 (1NC, 2C)	
	Only adapt players until they reached the same level	2 (2C)	
	reward with explicit adaptation	1 (1NC)	
Advantages of dynamic difficulty adaptation	Close games are more fun	8 (3NC, 5C)	22 (10NC, 12C)
	Gradual difficulty level	7 (4NC, 3C)	
	Avoid failing often or badly	3 (2NC, 1C)	
	Casual players won't give up easily	2 (2C)	
	Pushes players like a trainer	1 (1NC)	
	Avoid boredom	1 (1C)	
Dependencies for type of adaptation	Player's personality	5 (3NC, 2C)	11 (4NC, 7C)
	Mindset of players	4 (4C)	
	Type of game	2 (1NC, 1C)	
VR Requirements	VR can create a different reality for each player	7 (7NC)	8 (8NC)
	Virtual skills don't translate to real life	1 (1NC)	
Direction and receiver of adaptation	Increase for less skilled players	1 (1C)	3 (1NC, 2C)
	Bi-directional	1 (1NC)	
	Depends on skill difference	1 (1C)	
Proposed solutions for balancing mismatching skills in exergames	Static difficulty adaptation	11 (9NC, 2C)	23 (15NC, 8C)
	Co-op	8 (4NC, 4C)	
	Dynamic difficulty adaptation	4 (2NC, 2C)	
Examples of dynamic difficulty adaptation	Giving hints	8 (8C)	27 (5NC, 22C)
	Handicaps for more skilled player	5 (1NC, 4C)	
	Elo ranking	4 (4C)	
	Manipulating game elements (size, speed, distance of targets, allowance for mistakes)	4 (2NC, 2C)	
	Mario Kart is a good example	3 (1NC, 2C)	
	Catch-up system when falling behind	1 (1C)	
	Measure heart-rate to determine effort	1 (1NC)	
	Reward different amounts of points depending on level	1 (1C)	

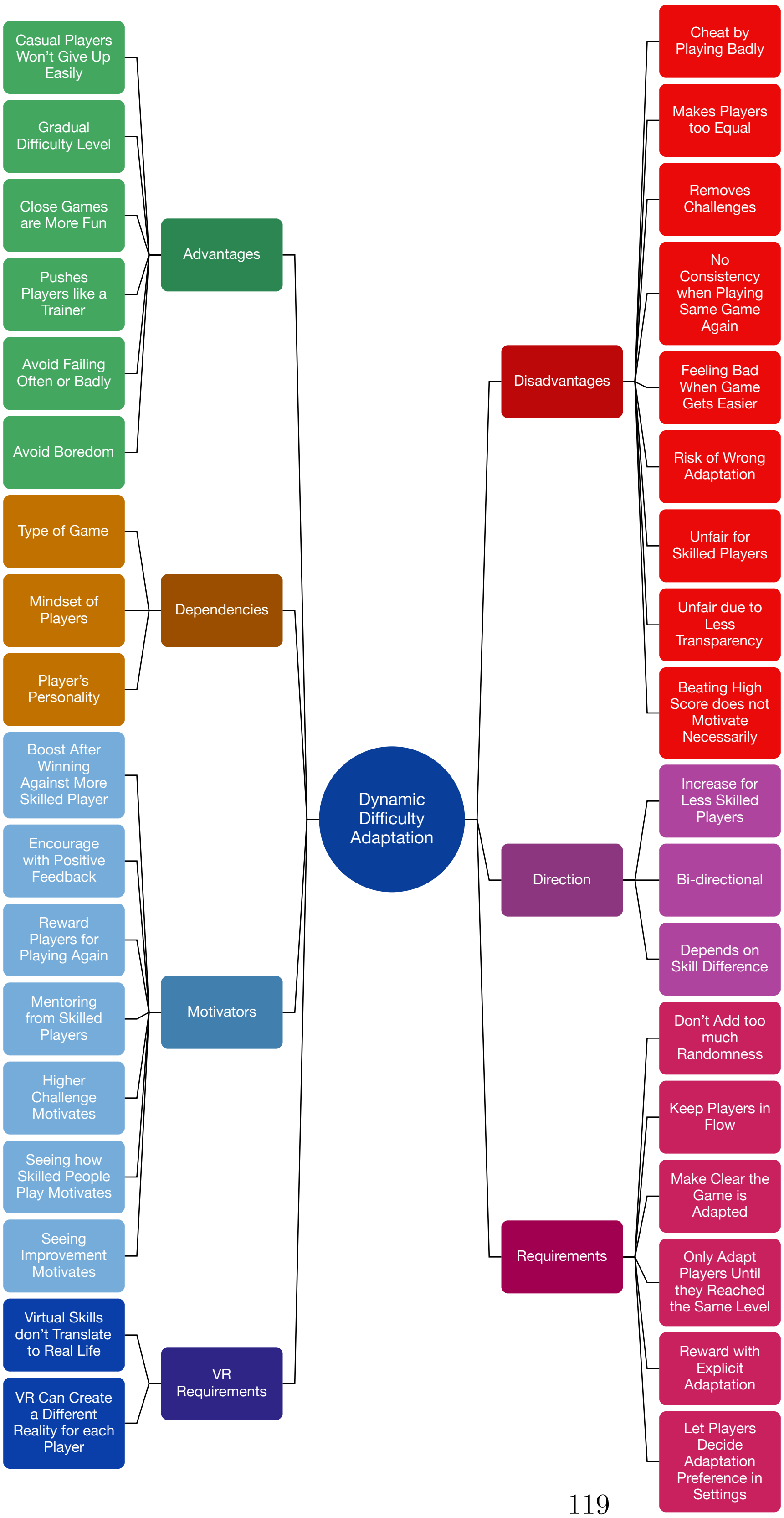
C.2 Mind-map Non-Competitive



C.3 Mind-map Competitive



C.4 Mind-map All Categories



D Design Focus Groups

D.1 Handout Sheet Stating Requirements

Requirements for dynamic difficulty adaptation

Task: Adapt an exergame to dynamically adapt the difficulty during the game. The exergame is a rhythm game similar to Beat Saber.

- **Create close games**
  If players fall behind, they can catch up. The player on the first place could lose their position again.

- **Avoid demotivation of the less skilled player**
  Less skilled players could feel demotivated when seeing that their game got easier because of poor performance. This can be avoided by either using less obvious adaptations or giving the player the choice of using an advantage in the game. Think of Mario Kart, where items have to be collected and activated to be used.

- **Make improvement visible**
  Seeing improvement is often more motivating than winning. However, losing all the time can be frustrating, too. Players also feel rewarded when the game gets more difficult due to their good performance -> use visible adaptation.

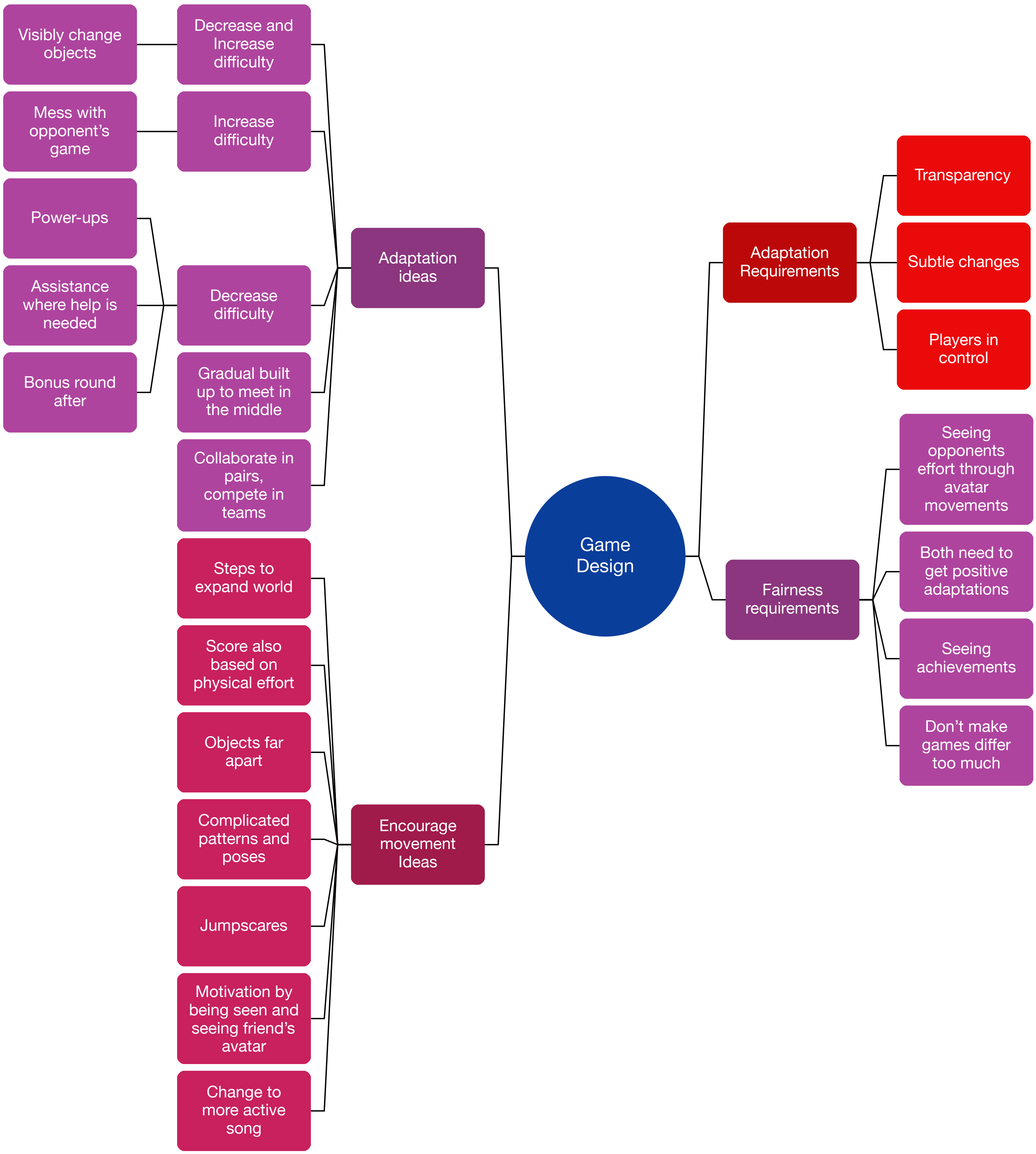
- **Add the right amount of randomness**
  Randomness can help to create a close game. Making it too random can be frustrating when putting effort into playing.

- **Use the advantage of VR: every player can have their own reality**
  The designer of the game can decide what players can see of other players: Can they see that their opponents game got adapted? Should they know how difficult or easy the opponent's game is?

D.2 Codes Sorted Into Categories – Focus Group 1

Category	Code	# Statements	Notes
Adaptation ideas	Visibly change objects	3	When asked what idea is their favorite, power-ups were clearly in favor. Messing with opponent's game is liked for more interactivity between players. 1 P. likes the gradual built up of difficulty until players meet in the middle and 1 likes collaboration as best idea.
	Mess with opponent's game	3	
	Power-ups	10	
	Assistance where help is needed	1	
	Bonus round after	1	
	Gradual built up to meet in the middle	3	
	Collaborate in pairs, compete in teams	3	
Adaptation requirements	Transparency	7	Transparency and adapting games in a way that they do not differ extremely is necessary to create a fair game. Transparency means that players can see how their and the opponents game adapts. Players also want to be in control as in activating an advantage/power-up.
	Subtle changes	3	
	Players should be in control	3	
Fairness requirements	Seeing opponents effort through avatar movements	4	Participants would like to see their friend as avatar, similarly they want to interact with them otherwise it does not feel like they are playing with friends. Achievements can be shown both within and after the game. Seeing movements of your friend can motivate either by being impressed of well-performed movements or by seeing that the friend put physical effort into playing.
	Both need to get positive adaptations	2	
	Seeing achievements	3	
Ideas to encourage movement	Steps to expand world	2	Complicated poses were liked because it adds a funny part to the game as in seeing your friend's avatar twisted. This idea also ties with the idea of placing objects far apart. Jumpscares were not liked. Score based on physical effort was liked and adds the possibility of catching up even if the player is not good in hitting targets. It also encourages more movement.
	Score also based on physical effort	1	
	Objects far apart	5	
	Complicated patterns and poses	8	
	Jumpscares	5	
	Change to more active song	1	

D.3 Mind-map Game Design Derived from Focus Group 1



D.4 Codes Sorted Into Categories – Focus Group 2

Category	Code	# Statements	Notes
Favorite idea	Combine power-ups and messing with opponent	10	Overall, participants would combine two ideas but the idea “messing with opponents game” was their favorite. Players have to activate the power-ups. They can do so by moving generally more during gameplay and filling up a loading bar that activates a power-up. The bar can load quicker for the player who is behind.
	Incentivize physical activity	6	
Requirements	Self-explanatory design	1	
	Do not mess too much with opponent	3	
	Player needs to put effort to get power-up	4	
	Provide a way to catch up when falling behind	2	
Ideas to activate power-up	Loading bar fills up with more movement and unlocks power-up	3	Loading bar is preferred and can be combined with incentivizing physical activity during the game.
	Make poses to unlock power-up	1	
Reasons for choice	Automatically change size of objects to change difficulty seems unfair because it is more about coordination skills	1	Participants liked “messing with oppoinents game” because it is interactive, fun and original.
	Favorite idea is fair because weaker player gets an alternative way to win the game which requires less skills but incentivizes physical activity	1	
	Likes interaction between players	5	
	Likes the game mechanic to mess with the other opponent because it is fun	5	
	Stronger player gets challenged	1	
Additional features	Add strategy to the game so that players have to decide when it is best to activate a power-up	1	Feedback about the current place within the game has to be incorporated to make players aware of how far behind/ahead they are.
	Calorie counter	2	
	Get bonus for streaks – hitting objects in a row without a miss	3	
	Give hints how to catch up during the game when falling behind so player knows what to do	1	
	Provide feedback during the game who is leading and who is behind	1	

E User Study

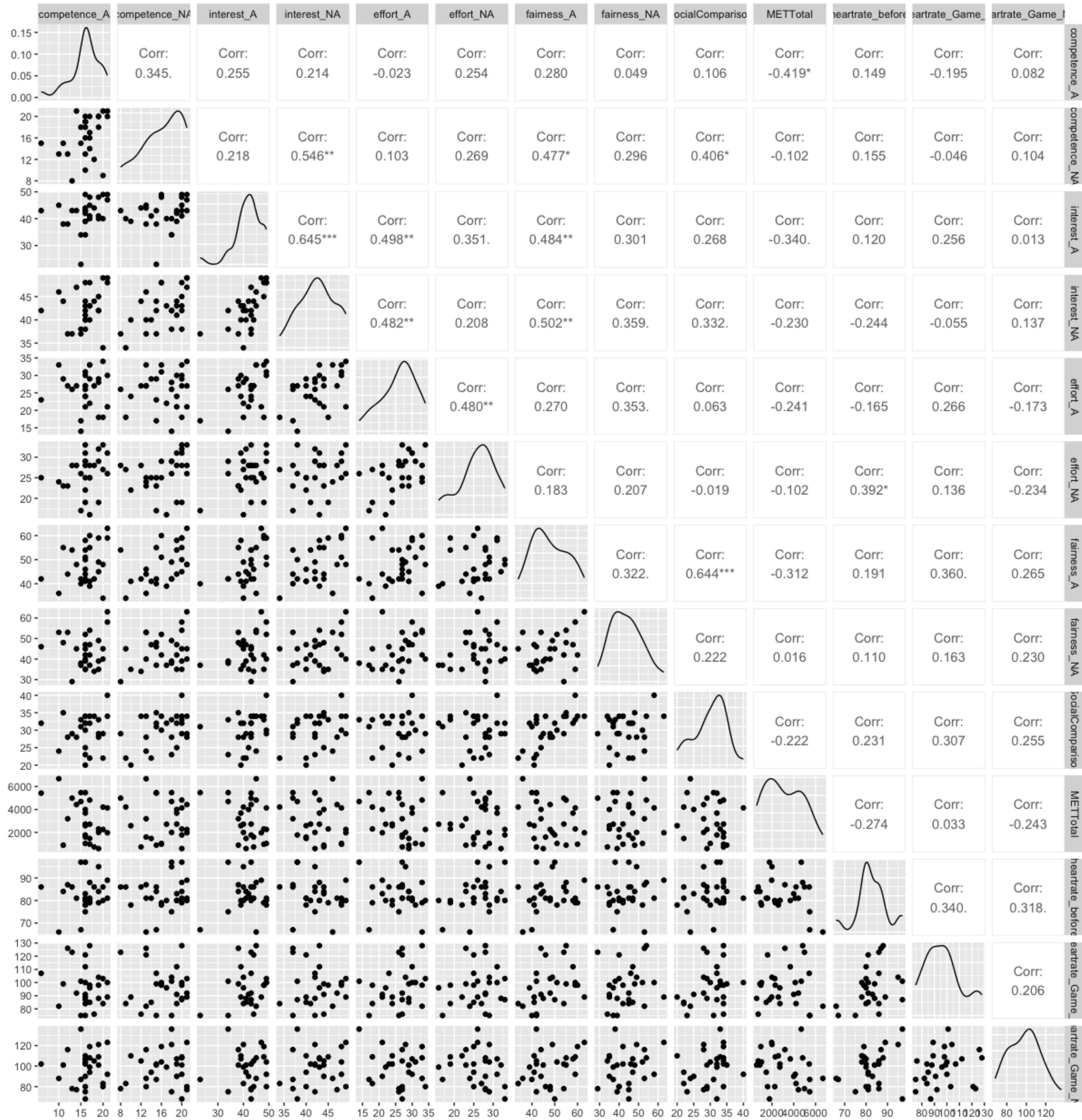
E.1 Codes Sorted Into Categories

Category	Code	# Statements	Notes
Favourite game	With Adaptation	20	
	No Adaptation	8	
	Likes both	6	
Advantages of adaptation	Fairly balanced game	28	
	Higher motivation	9	
	Achievement through earning bonuses instead of getting them for free	5	
	Increases challenges for skilled player	10	
	Many bonuses feel good	5	
	Makes game more intelligent	2	
Dislikes of adaptation	Spawning bonus blocks made the game more difficult	21	
	Attacks were too extreme and frustrating	14	
	Did not notice benefits from it	3	
	Game mechanics unclear	15	
	Wants to measure true skills	5	
	Jealous of bonuses	1	
	Unfair for skilled player	1	
Best context for adaptation	Good for casual games	3	
	Good for major skill differences	2	
	Adaptation only at the start, later without	1	
General improvements for game	Visual cues to see that opponents got hit by attack	1	
	Feedback and tips after game	2	
	Scoring hard to understand	2	
	Make adaptation full part of the game, not a mode to be chosen	2	
	Attacks are good addition for more interactivity between players (keep)	3	
	Rhythm did not match blocks	1	
	Make game faster	1	
	When score differ a lot, immediately decrease difficulty instead of offering possibilities	2	
Other notes	Too focused on game to notice adaptation	2	
	Surprised that player could suddenly catch up	1	

F Prototype Repository

https://github.com/mantzta/danceandattack

F.1 Correlations Across Variables



G User Study Survey Link

https://survey.uu.nl/jfe/form/SV_2f8osdtIG2AKtsa

H Ethics and Privacy Quick Scan

Response Summary:

Section 1. Research projects involving human participants

P1. Does your project involve human participants? This includes for example use of observation, (online) surveys, interviews, tests, focus groups, and workshops where human participants provide information or data to inform the research. If you are only using existing data sets or publicly available data (e.g. from Twitter, Reddit) without directly recruiting participants, please answer no.

- Yes

Recruitment

P2. Does your project involve participants younger than 18 years of age?

- No

P3. Does your project involve participants with learning or communication difficulties of a severity that may impact their ability to provide informed consent?

- No

P4. Is your project likely to involve participants engaging in illegal activities?

- No

P5. Does your project involve patients?

- No

P6. Does your project involve participants belonging to a vulnerable group, other than those listed above?

- No

P8. Does your project involve participants with whom you have, or are likely to have, a working or professional relationship: for instance, staff or students of the university, professional colleagues, or clients?

- Yes

P9. Is it made clear to potential participants that not participating will in no way impact them (e.g. it will not directly impact their grade in a class)?

- Yes

Informed consent

PC1. Do you have set procedures that you will use for obtaining informed consent from all participants, including (where appropriate) parental consent for children or consent from legally authorized representatives? (See suggestions for information sheets and consent forms on the website.)

- Yes

PC2. Will you tell participants that their participation is voluntary?

- Yes

PC3. Will you obtain explicit consent for participation?

- Yes

PC4. Will you obtain explicit consent for any sensor readings, eye tracking, photos, audio, and/or video recordings?

- Yes

PC5. Will you tell participants that they may withdraw from the research at any time and for any reason?

- Yes

PC6. Will you give potential participants time to consider participation?

- Yes

PC7. Will you provide participants with an opportunity to ask questions about the research before consenting to take part (e.g. by providing your contact details)?

- Yes

PC8. Does your project involve concealment or deliberate misleading of participants?

- No

Section 2. Data protection, handling, and storage

The General Data Protection Regulation imposes several obligations for the use of **personal data** (defined as any information relating to an identified or identifiable living person) or including the use of personal data in research.

D1. Are you gathering or using personal data (defined as any information relating to an identified or identifiable living person)?

- No

Section 3. Research that may cause harm

Research may cause harm to participants, researchers, the university, or society. This includes when technology has dual-use, and you investigate an innocent use, but your results could be used by others in a harmful way. If you are unsure regarding possible harm to the university or society, please discuss your concerns with the Research Support Office.

H1. Does your project give rise to a realistic risk to the national security of any country?

- No

H2. Does your project give rise to a realistic risk of aiding human rights abuses in any country?

- No

H3. Does your project (and its data) give rise to a realistic risk of damaging the University's reputation? (E.g., bad press coverage, public protest.)

- No

H4. Does your project (and in particular its data) give rise to an increased risk of attack (cyber- or otherwise) against the University? (E.g., from pressure groups.)

- No

H5. Is the data likely to contain material that is indecent, offensive, defamatory, threatening, discriminatory, or extremist?

- No

H6. Does your project give rise to a realistic risk of harm to the researchers?

- No

H7. Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?

- Yes

H8. Is there a realistic risk of any participant experiencing a detriment to their interests as a result of participation?

- No

H9. Is there a realistic risk of other types of negative externalities?

- No

Ethics Warning. **As you replied yes to one (or more) of H1-H9, a fuller ethical review is required.**
Please provide more detail here on the potential harm, and how you will minimize risk and impact:
In the study participants will play a virtual reality game. Therefore, it is possible that participants may experience discomfort as in motion sickness. They will be informed about this risk and also that they should stop playing if feeling uncomfortable. Additionally, they will be informed that they can stop playing or participating at any time.

Section 4. Conflicts of interest

C1. Is there any potential conflict of interest (e.g. between research funder and researchers or participants and researchers) that may potentially affect the research outcome or the dissemination of research findings?

- No

C2. Is there a direct hierarchical relationship between researchers and participants?

- No

Section 5. Your information.

This last section collects data about you and your project so that we can register that you completed the Ethics and Privacy Quick Scan, sent you (and your supervisor/course coordinator) a summary of what you filled out, and follow up where a fuller ethics review and/or privacy assessment is needed. For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information. Please see the guidance on the ICS Ethics and Privacy website on what happens on submission.

Z0. Which is your main department?

- Information and Computing Science

Z1. Your full name:
Tamara Mantz

Z2. Your email address:
t.m.mantz@students.uu.nl

Z3. In what context will you conduct this research?

- As a student for my master thesis, supervised by::
  Hanna Hauptmann

Z5. Master programme for which you are doing the thesis

- Human-Computer Interaction

Z6. Email of the course coordinator or supervisor (so that we can inform them that you filled this out and provide them with a summary):
h.j.hauptmann@uu.nl

Z7. Email of the moderator (as provided by the coordinator of your thesis project):
graduation.hci@uu.nl

Z8. Title of the research project/study for which you filled out this Quick Scan:
Balancing Multiplayer Exergames in Virtual Reality with Dynamic Difficulty Adaptation

Z9. Summary of what you intend to investigate and how you will investigate this (200 words max):
I will investigate the effects on motivation/fairness/competence when playing multiplayer exergames in virtual reality (VR) with difficulty, that dynamically adapts during gameplay. For this, I will conduct focus groups and interviews to gather insights about how to design difficulty adaptation. Afterwards, I will implement a prototype of a game that includes difficulty adaptation. In the study, participant pairs will play against each other in two conditions: A. no difficulty adaptation present B. difficulty adaptation is present. After each condition both participants are asked to answer questionnaires regarding their perceived motivation/fairness/competence. Additionally, I will conduct a short interview with the participants to understand what they liked and did not like and why. At the end, they are asked to fill out a questionnaire about demographic data, their weekly physical activity and how familiar they are with VR and exergames.

Z10. In case you encountered warnings in the survey, does supervisor already have ethical approval for a research line that fully covers your project?

- No

Scoring

- Privacy: 0
- Ethics: 1
