Assessing the Effect of Gamification on various aspects of Sport Exercises through an Augmented Reality Exertion Interface

Nynke Zwart
Radboud University Nijmegen, 6525 AJ, Nijmegen
http://www.ru.nl/english/
n.zwart@student.science.ru.nl

Abstract. This research focuses on assessing the effects of gamification on sports exercise. An Augmented Reality (AR) approach was used to research this. Since AR is growing in popularity, it is new and interesting to perform this research in an AR context. A baseline exercise program (BEP) and gamified exercise program (GEP) were made for two randomly assigned participant groups. It has been found that scores on game flow and the duration of workout have significantly higher results for the GEP group. Difference in mood turned out not to be significant between the two groups. The research has great potential for further improvement.

Keywords: Sports · Gamification · Augmented Reality.

1 Introduction

The focus of this research lies on assessing the effects of gamification on sports exercises. Finding motivation and joy in working out can be difficult at times, as it can depend on one’s own feelings and situational factors [14]. This lead to the idea whether game design could be used to make working out more motivating and fun [4]. Thus, the application needed to include an exertion interface: an interface that requires physical effort from the user [16]. While searching for a technique that could gamify the sports experience, yet would not hinder the performance either, an augmented reality approach was chosen. This approach of gamification and AR combined, as explained in section 1.1, is an unexplored field of research as well.

In augmented reality (AR) interfaces, digital or computer generated information is overlayed over a real-time environment [11]. This information can contain images, audio, video, touch or haptic sensations, amongst others. Head-mounted or handheld displays (HMDs) are typically used to make this information appear in the user’s environment [3]. However, AR is not limited HMDs, even though authors often choose to include HMD’s in their definition of it. Mekni and Lemieux observed this in their paper about AR applications, challenges and future trends [15]. For example, Anderson et al.’s research demonstrate a non-HMD example:
an AR Mirror System [1]. This system was created by using half-silvered mirror film and diffuse film. By using a projector, participants could see themselves and virtual objects in real time on this mirror. Since indoor sports often take place in front of a mirror, this set-up inspired my research approach. Hence, for my application, virtual information is displayed live on a regular screen that mirrors the real environment by use of a webcam. To track the movements of the user and estimate where virtual objects should be placed on the screen, a Kinect sensor is also used. This movement tracking will provide the Artificial Intelligence (AI) component in my research.

1.1 Related work
Gamification of sports is already a well explored field [16]. One might think of smartwatches, online communities and various fitness applications. Larsson put some of these already existing applications to the test in his research [13]. The participants in his research wanted to get in better shape than they currently were and several gamification services were tested to see if they could change this. Most participants argued that because of external rewards such as visualised data about their exercises and achievement systems, they felt more motivated to keep on training. Whether their training sessions were of longer duration, was not mentioned. The participants also felt sure that if the external rewards would stop, they would not continue the service. Giannakis et al.’s research on sport services also brought to light that while services can promote working out, there is a trade-off of increased anxiety and disorientation [8]. According to the research, sports software should be modest about how much and what type of visual information it provides to the user. Too much information can overwhelm the user and cause stress. Hence, the interaction between system and user needs to be straightforward and not overly complicated.

There are also studies on gamification of sports that research self-made applications, such as Mueller et al’s exertion interface applications [17]. As mentioned before, exertion interfaces require physical effort from the user [16]. It is a unique kind of interface and differs from the other services in the way that it deliberately requires physical activity. A smartwatch, for example, will not force a person to exercise, but the exertion interfaces that Mueller has developed over the years, do. Mueller has created multiple applications to connect two people doing sports together over a distance. He has managed to create exertion interfaces for sports as soccer and airhockey. They work as follows: the person at a distance is projected onto a wall. The actions of each person affect the other person, so there is true interaction between them. They reported a significant difference for fun in their research on the soccer interface [17]. The exertion interface group, as opposed to the keyboard interface group, enjoyed playing much more. Fun is a factor that will also be looked at in my research.

Another well explored field is that of AR games. To name one example out of many: “ARQuake”. In this game created by Piekarski and Thomas, mobile users fight virtual enemies in a real environment [21]. They used a HDM for this. However, it is easy to imagine playing this game on your mobile phone as well, which
greatly expands the target audience. In 2005, one of the earliest AR games for a phone was made. The game, AR tennis, was played on Nokia mobile phones and was a collaboration between two players. It was developed by Henrysson et al [9].

Now, with the rise of personal mobile devices capable of producing AR environments through mobile applications, the vast potential of AR has begun to be explored in an easy accessible environment [15]. Newer phones are faster and have grown in screen size, which is advantageous for AR games. Over the years, many others have created games that use a mobile devices as their basis. Arguably the most popular AR game to have ever been developed is Pokémon Go. Pokémon Go motivated players who were previously sedentary to become more physically active. If they did not play Pokémon Go, they were likely to walk about 110 minutes less a week [26]. It proves how AR games can be used as a starting point to help people get more active.

Now that examples of gamification of sports and AR games have been mentioned, it is noticeable that not much has been said about the combination of these two fields. The reason behind this is that this combination is still quite unexplored. One of the few projects that take up this combination of fields is the research about the YouMove system by Anderson et al [1]. The YouMove system helps users to learn physical movement sequences. It does so by usage of a Kinect sensor and an Augmented Reality Mirror system. The user sees himself/herself projected on the mirror system in real time. The target movement is overlayed on this projection. By use of shapes and colours, a target skeleton is made that users can follow to learn movements precisely. Besides that, the user gets scores based on the similarity between their movement and the target movement they have to learn. Their movement is then compared to the target movement data, which is how the similarity score is calculated. There were two group conditions: the YouMove condition and the Video condition. It was found that the YouMove system improved learning, as the retention scores were higher than those for the Video condition.

Aside from Anderson’s research, not many papers on AR gamification of sports can be found. With this thesis, I address this gap in research for various reasons. For example, the YouMove system is focused on educating its participants in learning physical movement sequences [1], whereas I focus on measuring the motivation and enjoyment aspects of exercise instead of learning. This is a new direction which can be a valuable addition to the YouMove research. Besides that, the use of AR is predicted to grow steadily over the coming years, so it is important to latch onto it and provide more research about its potential in multiple fields. While this Related Work section brings forward some interesting research that has been done already, I will mention more relevant research in section 2 about my BepNGep application. This section will explore more design related issues. Lastly, in section 3 on Methods and Materials, there is more information on how I plan to obtain my measurements by using works from Pelletier (Sports Motivaton Scale [20]), Jones (Sport Emotion Questionnaire [10]) and Sweetser and Wyeth (GameFlow criteria [25]).
1.2 Research questions and hypotheses

My research questions are as follows:

1. Can differences be found between the baseline and game group in relation to joy, score on game flow of the program and time of exercising? What differences can be seen when dividing the participants in personal motivation groups?

2. How does the time of exercising depend on joy and game flow score?

For question 1, I expect to find that the game group scores significantly higher on all variables than the baseline group. The baseline group will namely play a version of the exercise program that is very basic and contains minimal information. The results for this group will thus reflect how they feel after a usual workout. However, the game group is exposed to a program containing multiple stimuli in the form of scores, sounds and more. Game applications that provide rewards for accomplishments have immediate effect [19], so I expect participants to want to push further. When the program is captivating enough, participants could become less aware of their surroundings and feel emotionally involved in the game [25]. I believe these factors will then result in a higher exercising time that reflects being motivated. For the game group, I expect to find a higher score on joy. This will be brought about by the design of the game, that will put exercising in a new light and give them a novel experience. The addition of game elements to the workout are expected to higher the score of the game group on game flow as well. Game flow reflects the players enjoyment: this score is based on game criteria that are part of the GameFlow model by Sweetser and Wyeth’s (see section 2.1). If no significant results are found, it would mean that the participant gets as much joy from the baseline as from the game program. I strive to create a game that is sufficiently different from the baseline program and thus do not expect this to happen. If the game and baseline group can be split into equally sized groups of extrinsically and intrinsically motivated participants, I expect that the same significant results will be found, but that the means will differ. Since my game program will mainly focus on providing external rewards, I expect extrinsically motivated participants to play longer, rate game flow higher and have more fun than intrinsically motivated participants. This is because extrinsically motivated people are more motivated by external rewards (see definitions in section 2.1).

Concerning the second research question, I expect to find a correlation between duration of the workout, game flow and joy as follows: the higher the game flow score and score on joy of the participants, the longer the duration of workout. When people perform sports it is often because they are motivated and/or enjoy it, which means they are willing to invest time in it.

2 BepNGep

For this research, two exercise programs were created, a baseline exercise program (BEP) and gamified exercise program (GEP). Several design considerations
took place before the final design was completed (see section 2.2) and the study could be conducted (see section 3.2). This process is explained below.

2.1 Design considerations

My research questions require good reasoning about how to shape my applications. Or to summarise it: how to design for motivation and enjoyment? It turns out there are many opinions about this. I will first dive into motivation: what is it exactly?

**Internal and external motivation** The research on motivation in gamification can be categorised into two categories: one that advocates the work on the user’s external (extrinsic) motivation and one that focuses on internal (intrinsic) motivation [13]. Internal motivation refers to “doing something because it is inherently interesting or enjoyable” [22]. It is often compared to external motivation as its opposite: “doing something because it leads to a separable outcome” [22]. These outcomes are often referred to as rewards. External rewards are widely used to gamify a certain task. It’s an effective method to control people’s behaviour [5]. While these rewards are regularly used by researchers to increase intrinsic motivation, they mainly enhance extrinsic motivation [12]. According to Deci, Koestner and Ryan, extrinsic motivations can even reduce the initial intrinsic motivation. In other words, reward contingencies undermine people’s taking responsibility for motivating or regulating themselves [5]. This is why gamification must be handled with careful consideration for the individual. A different side of the coin is shown by Zichermann and Cunningham, who claim that gamification can be used to provide external motivation which will help a person find (increased) internal motivation. For this to happen, the user needs to stay interested in the external rewards of the system until the internal motivation is high enough for him/her to perform the task out on their own [27]. Nicholson’s research about Meaningful Gamification (read more in section 2.1) addresses how to design for a long game path like this.

It is clear that motivating a person is a delicate thing, since the effects of the chosen approach can differ from person to person, depending on how internally motivated they are before the task. My application will primarily be built on external rewards. If the participant group can be split into externally and internally motivated people, it is interesting to see if differences can be found in their mood and rating of the program.

**Designing for motivation** The designer of a game has the controls over which behaviours are rewarded [19]. There are many frameworks, models and heuristics developed that should guide a designer towards the right approach. I will discuss the most prominent ones that guided my final design.

**BLAP Gamification** BLAP gamification, a term defined by Nicholson, is the concept of adding Badges, Levels/Leaderboards, Achievements and Points
to a real-world setting [19]. BLAP gamification systems are relatively easy to implement and have an immediate effect on the user. This is because the instant rewards engage people and play on their extrinsic behaviour. For short-term change, a BLAP system is suitable. Often, in research where BLAP gamification is used, the long-term effects are rarely discussed. My research will focus mainly on BLAP elements, yet also includes the autonomy component of the self-determination theory (see section 2.1). To obtain a broad view of design possibilities, this theory and other elements that are said to higher internal motivation are mentioned as next.

**Self Determination Theory**
The self-determination theory is a macro theory of human motivation concerning people's inherent growth tendencies and innate psychological needs. It focuses on intrinsic motivation and how to maintain it. They state that it is a combination of three psychological needs: autonomy, competence and relation [6]. Autonomy refers to the sense of will when performing a task. Autonomy is high when activities are performed out of personal interest. It is shown that intrinsic motivation of individuals can be improved by providing opportunities to choose, which higher the perceived autonomy [23]. Instructions should not be too constraining and feedback should support the user's exploration. In my program, users are completely in control of how long they want to perform each exercise and when they wish to go to the next exercise or stop the program. This can be seen as a sense of autonomy. The second need, competence, is the need of people to participate in challenges and feel competent and efficient. Opportunities for acquiring new knowledge or skills need to be abundant and either challenge the user or encourage the user by giving positive feedback. This improves the perceived level of competence and therefore intrinsic motivation. Lastly, relation refers to the need to feel connected to others. The integration of a social network within a game is a mechanism that can reinforce motivation [2].

**Meaningful Gamification**
The concept of meaningful gamification is that it focuses on using game design elements to help users find a meaningful connection to the real-world setting [19]. Nicholson created this framework, which considers six alternative ways of game design elements as opposed to BLAP gamification [7]. They are summerized below:

- **Reflection**: creating situations where users can reflect as to discover personal connections with the real-world setting. If the connection to “others” in the “Relation” component of the STD theory is read as a connection to the real-world setting, the “Relation” component is relevant here.
- **Exposition**: using narrative and user created stories in order to create deeper connections to the real-world setting. This refers back to “Relation” in the STD theory, for the same reason as the Reflection element.
- **Choice**: allowing the user to select paths and develop goals within the real-world setting that feel more meaningful to them. “Autonomy” is relevant here: the user picks what he/she wants to work on.
– *Information:* providing the user with information between the gamification activities and the real-world setting. This refers to “Competence” in the STD theory.

– *Play:* creating a space where the user can choose how he/she wishes to engage with different gamification activities in the real-world setting. “Autonomy” is relevant here.

– *Engagement:* using gamification to connect users to a community of practice that surrounds the real-world setting. This element falls back on “Relation” in the STD theory.

**Combining gamification concepts** There is a way to combine BLAP gamification and meaningful gamification. A game can start with rewards, but slowly introduce meaningful connections to the real world. The hope is that eventually, the user will find this connection and no longer requires rewards. An example gamification system is shown in figure 1 [7]. The loop reflects BLAP gamification. If short term change in motivation is the goal, this loop can be followed. Yet, if the purpose is long-term change, rewards by BLAP gamification will probably not be enough to maintain interest and could harm intrinsic motivation [5]. More real-world information and communities should be implemented in the game: the user needs to be guided out of the loop. As the user becomes more of an expert, the system becomes less prominent and more of the real-world context comes forward.

![Fig. 1: The combination of BLAP and Meaningful Gamification.](image-url)
Designing for enjoyment  In Ryan and Deci’s definition of intrinsic motivation, the word “joy” is mentioned. Motivation itself is defined as the reason(s) for acting a particular way. This does not automatically imply that there is emotion involved. However, emotions are rooted deeply within ourselves and often drive us when it comes to certain tasks. Mueller stated that the concept of fun as a motivational factor is a growing area of inquiry in the HCI domain, hence it is interesting to take a look at, which I’ll do via the concept of Flow [17]. Flow is a widely accepted model of enjoyment and will be explained in the following section.

Game Flow  Flow is a widely accepted model of enjoyment. GameFlow, the model Sweetser and Wyeth created, focuses specifically on games. It contains eight elements: concentration, challenge, skills, control, clear goals, feedback, immersion and social interaction. Each of these elements includes a set of criteria for achieving enjoyment in games.

Flow is based on the assumption that elements of enjoyment are universal and proposes a model that summarises concepts that are common to everyone when they experience enjoyment. Most flow experiences occur with activities that are goal-directed, bounded by rules and that require mental energy and the appropriate skills [18]. It is clear to see that STD theory elements “Competence” and “Autonomy” are related to a flow experience.

Sweetser and Wyeth analysed heuristics in games literature that overlap closely with the elements as flow and as a result came up with the GameFlow model [25]. This model is especially meant to assess player enjoyment in games. It consists of eight core elements, namely: concentration, challenge, skills, control, clear goals, feedback, immersion and social interaction. These elements are also summarised in figure 2 under the header “Games Literature”. “The Game” is not included in this count of eight elements, because it is the game itself. “Social Interaction” is added as a new element compared to flow literature, since user-experience game literature often features this concept. It adds the “Relation” aspect of the STD theory. All other GameFlow elements are closely interrelated and interdependent to the flow elements [25]. Central criteria for each element were developed as well and guided me in creating my questionnaire to assess how game flow is experienced in my own programmes.

Furthermore, I believe that per element, a design choice can be made between using meaningful gamification or BLAP gamification. For example, Nakamura et al. state that the flow experience is intrinsically rewarding, such that often, the goal is “an excuse” for the process [18]. The use of the word “often” does not rule out that external rewards can bring about flow in people. Other elements like challenge can be designed by carefully selecting when certain behaviours are rewarded (BLAP gamification) or giving the user the freedom to figure out themselves what the correct behaviour is (meaningful gamification). Both can bring a user joy, if carefully thought out. These are just a few examples of reasoning on how to achieve a flow state for users, to demonstrate that multiple
ways are possible. These GameFlow criteria are thus a good option to use when measuring the flow people feel while using my exercise program.

<table>
<thead>
<tr>
<th>Games Literature</th>
<th>Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Game</td>
<td>A task that can be completed</td>
</tr>
<tr>
<td>Concentration</td>
<td>Ability to concentrate on the task</td>
</tr>
<tr>
<td>Challenge</td>
<td>Perceived skills should match challenges and both must exceed a certain threshold</td>
</tr>
<tr>
<td>Player Skills</td>
<td>The task has clear goals</td>
</tr>
<tr>
<td>Control</td>
<td>Allowed to exercise a sense of control over actions</td>
</tr>
<tr>
<td>Clear goals</td>
<td>The task provides immediate feedback</td>
</tr>
<tr>
<td>Feedback</td>
<td>Deep but effortless involvement, reduced concern for self and sense of time</td>
</tr>
<tr>
<td>Immersion</td>
<td></td>
</tr>
<tr>
<td>Social Interaction</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Fig. 2:** Mapping of elements from games literature to flow elements [25].

**Linking motivation and joy to game elements** Gamification is often supposed to be an effective instrument to foster motivation. However, research and investigations about the motivational pull of gamification from a psychological perspective is scarce [24]. In Sailer et al.’s research, several game elements are matched with motivational mechanisms. Their research is insightful as it provides information on how people can react to specific game elements. Below, I have itemized the game elements that I use that are also discussed in their research. For the full explanation about my program design, see section 2.2.

- **Points:** these very basic game elements can address motivational mechanisms that mostly refer to the behaviourist learning perspective. The provide rewards for executed actions and function as immediate positive reinforcements.
- **Badges:** these are visual representations of achievements. Multiple feelings can be attached to these badges. They can foster a player’s feel of competence (perspective of self-determination) or need for success and/or status (trait perspective). If it is known how badges can be obtained, they can also be seen as goal setting (cognitive perspective).
- **Quests:** these are small tasks. The rewards for completing these tasks are usually known. They provide clear goals, emphasize the importance of a player’s action within a situation and address motivational mechanisms connected to the cognitive perspective.

Other elements Sailor talked about in his research are Leaderboards, Progress bars or performance graphs, meaningful stories and avatars and profile development [24]. Due to time constraints, these elements were not implemented, but they reappear later in the Discussion section (see section 5).
2.2 Program Design

Below, the features of BEP and GEP are explained in more detail. A quick side by side comparison of the programs can be seen in figure 3. A video on the programmes can be watched through the link in my references [28]. By seeing the programs in motion, their flow will be clearer than by studying still images only.

![Fig. 3: Baseline exercise program and gamified exercise program comparison.](image)

**Baseline Exercise Program** This version of the program is very basic and gives minimal information. This minimal information consists of which exercise needs to be performed. This is written in text and visualised by a video/image that is displayed on the screen. The user also gets an explanation beforehand on how to use the program. Most importantly is the word “next”: users must say this aloud when they want to go to the next exercise. They can also stop anytime that they want, either entirely or resume later if they need a small break.

**Gamified Exercise Program** The gamified exercise program contains the same elements as the baseline exercise program and more. It is extended by addition of the following elements:

- **Points**: each exercise movement is coupled to collecting coins. Coins were picked because money is valued in society as a reward for hard work. The game coins are aligned to the body in such a way, that they also guide the movement. There is no time limit in collecting the coins. The user can choose how long he/she wants to spend on each movement. However, more points are earned with more iterations. The user can see his/her points on the right side of the screen.

- **Badges**: the point system is figured out easily. This is why an achievement system is included too, in the form of badges. For certain amounts of points, badges are earned. These badges appear in the left side of the screen.
– **Quests:** a small task appears several times throughout the program. It is an extra points system, where coins are worth more than the usual one point per coin collection. The quest duration is fixed.
– **Sounds:** when collecting a coin, obtaining a badge or exercising during the quest, sounds are used. They differ from each other to make it clear to the user what is happening and to give direct feedback so a user can focus more on the task at hand.
– **Text:** text is used to announce achievements and quests. It also indicates the score and name of exercise. During the explanation of the programme, users are informed with text of what they should do and how to control the program.

![Fig. 4: A screenshot made while using the GEP program.](image)

Figure 4 shows a screenshot of the GEP program on which some of the above named elements can be seen. Lastly, the chosen exercises for the gamified exercise program are also part of the design and will be talked about in the following section.

**Chosen exercises** Careful consideration has been put into which sport exercises could benefit from an AR system that uses this set-up. The user will experience the program the best when he/she faces the screen, which is why the exercises are based on this principle. All exercises are performed standing on two feet, unless mentioned otherwise. The possible exercises are summarized below and sorted in the order they appear during the program (see figure 8 in addition). The user cannot chose which one comes next, but does have the control to decide how long they wish to perform each exercise. Furthermore, it is made sure that other body parts cannot collect the coins beside those that are meant to collect them.
Fig. 5: Exercises of the final GEP programme.

Fig. 6: Detected joints by the Kinect V2.
- **Jumping Jacks:** The coins appear relative to the head and spine base. The hands have to collect the coins.
- **Side Lunge:** The coins appear relative to the feet and have to be collected by the feet as well.
- **Woodchops:** The coins appear relative to the middle of the spine. The coins are collected by the hands. An extra weight may be used for this exercise.
- **Burpees:** The coins will appear relative to the middle of the spine and the feet. The coins have to be collected by the hands.
- **Standing Side Bend:** The coins appear relative to the middle of the spine. The hands need to collect the coins. An extra weight may be used for this exercise.
- **Power Side Kick:** The coins will appear relative to the spine base. The coins have to be collected by the feet.
- **Reverse Lunge Kick:** The coins will appear relative to the spine base. The coins have to be collected by the feet. This exercise only appeared in the pilot study and was later removed. For more explanation, see section 4.1.

3 Methods and Materials

In this section, the pilot study and final study will be explained in detail, starting with the pilot study.

3.1 Pilot Study

When a first design of the GEP program was made, it was tested by means of a pilot study. The progress of this program, compared to the BEP, was most crucial to test. Below, a summation can be found of the features of the test version, how the pilot was done and what the results were.

**Pilot implementation** The GEP program gave instructions about the following things:

- Start position, marked with an “X” on the ground.
- The fact that the participant will perform a set of sport exercises.
- How to go to the next exercise. This can be done by saying “next” in a loud and clear voice. The participant gets a chance to test this beforehand, too.
- What the workout interface looks like. On the right side of the screen, the score number is displayed. Here, it can also be seen which exercise needs to be performed. On the left side of the screen, badges appear, once they are earned.

Features which were not told to the user were as follows:

- The purpose of the game is to collect coins. By performing the exercise right, coins are collected. The exercises that were implemented appeared in the following order: Jumping Jacks, Power Side Kick, Side Lunge, Woodchops and Reverse Lunge Kick.
At that time, there was only one badge, which appeared when more than 50 points were collected. It is announced with text when a badge is achieved.

There was one moment in the game where coins were worth four times as much as usual. This phase was entered when a person reached more than 20 points and continued for 10 seconds. It is announced with text when this phase starts and marked with different coloured overlays as well as a change of colour for the scoring number.

Pilot experiment Two people took part in the pilot. In order to assess the completeness of the gamified program, the following questions were asked after performing the exercises.

- Rank the following features (on top = most favoured, on bottom = least favoured): score tracking, example video of exercise, sounds, collection of objects, voice commands, achievement badges, extra points mechanism.
- Which features could be improved?
- What features do you think were missing?

After these open questions, Questionnaire 3 from the appendix was filled in by the participants. This questionnaire takes two criteria from each of the eight important game elements, namely: concentration, challenge, player skills, control, clear goals, feedback, immersion and social interaction. To get the total score, all points are added up and divided by 16. The final score on game flow lies on a scale of 1 to 5, as is done as well in Sweetser’s research. Questions 2, 9 and 16 are phrased in an inverse manner, this needs to be taken into account when adding up all the points. The results that followed from this pilot are written down in section 4.1.

3.2 Final study

After the insights that were gained through the pilot were processed, the actual study was held. The study was performed in a room at Raboud University Nijmegen. For this research, a baseline and gamified AR application were created (see section 2.2), which meant there were two groups of participants. The baseline group played the baseline exercise program, the game group the gamified exercise program. Participants performed one of the conditions, only once. Thus, the design of the study is experimental and cross-sectional. The independent variable in this research is the randomly assigned group (qualitative). Game flow, joy and duration of workout are quantitative variables. 40 participants between the ages of 18 and 30 years (26 males, 14 females) took part in the study. Convenience sampling was used to find the participants.

The exercises were tailored so the participant always faced the screen (see section 2.2). The participant had control over how long he/she performed each exercise. Via speech recognition, he/she could command to go to the next exercise. The maximum amount of workout time was 10 minutes.
Before and after the exercise, questionnaires had to be filled in. To form questionnaires, the Sports Motivation Scale of Pelletier [20] was used, amongst others. This questionnaire by Pelletier assesses whether a person is intrinsically or extrinsically motivated to do sports (pre-measurement). The Sport Emotion Questionnaire by Jones [10] was used too: this questionnaire creates a score on different levels of emotion, namely: anxiety, dejection, excitement, anger and happiness. A pre- and post-measurement allowed to calculate a score on mood difference. For the post-measurement, participants had to indicate how they felt during the exercises, not after. A positive difference indicates that a participant became more joyful. Lastly, Sweetser and Wyeth’s research about GameFlow [25] steered the questionnaire to evaluate my application. Two main criteria from each of the eight elements were used to create my own questionnaire. All the questionnaires and other questions that were used during my experiment can be found in the appendix. The questionnaires for these variables were made in such a way that a total score could be calculated as well to get quantitative variables. A section on the scoring system in the appendix explains how to calculate these scores.

Several analyses have been done in order to answer the research questions. For research question 1 (see section 1.2) independent samples t-test’s are used to evaluate each quantitative variable (joy, game flow, duration of exercise) compared to the between-subjectfactor Group (BEP or GEP). A Pearson correlation coefficient was computed to be able to answer research question 2, with joy, game flow and duration of exercise as variables.

Lastly, Unity, Vuforia and Kinect plugins were the necessary software compo-
nents for this research. A Kinect sensor, screen and webcam were the necessary hardware components. A yoga mat was used to make the exercises more comfortable and weights were available to optionally use during selected exercises (see figure 7 for the experimental set-up).

4 Results

4.1 Pilot study

The results of the pilot were as follows: when it came to the open questions, both participants rated score tracking and the extra points mechanism as very favourable, as it was in both their three most favourable features. After those features, the achievement badges were a favourite. BLAP gamification seems to be a good approach to achieve favourable features. One participant mentioned that she was very motivated to get a high score number, since she wanted to achieve more badges after she got the 50+ point badge. The example video was rated low to lowest. During exercising, it was really useful to get going, but it is not exciting. The opinions about the voice commands differed from a last place to a second place. The collection of objects and sounds were given middle positions.

Several improvements were mentioned. For example, the Reverse Lunge Kick and Power Side Kick did not always respond well to the participant’s movements. This is partly due to the Kinect, which does not always recognize the joints in a correct way. When a foot is lifted high, it can recognize this limb as a hand for example. This way, the coin cannot be collected because it will only react to feet during these exercises, to avoid cheating. The failure of the Kinect to recognize the feet could also be due to the backdrop. The objects that stood on the floor behind the subject could have interfered with recognizing the feet in a way that was no problem for hand recognition. This will be looked into, as I wish to avoid frustration for the participants.

Furthermore, the longer the participant exercised, the more the webcam was lagging. It is important for the webcam image to be precise and change in real-time. The Kinect has no problem with lagging in skeleton tracking during the program. A different webcam is used than the camera in the Kinect due to a Unity’s incapability to recognize the Kinect camera. An idea has already been formed about how to improve the lagging webcam image.

One participant was afraid that after a while, sounds could start to get annoying. During the actual experiment, I want to put music on in the background so there are other sounds a participant can focus on. Furthermore, different sounds are used for obtaining badges, the extra points mechanism and the coin collection. But even for a specific sound such as for coin collection, different variants could be used.

Sometimes, participants forgot to say “next” and needed to be reminded they could use this, even though it was explained to them. The keyword recognizer in Unity does not work very well either, so I had to manually press the down arrow on my keyboard to call upon the next exercise a few times. Saying “next” is
simply a handy mechanism for the participant to indicate they want to perform another exercise. However it is handled, is not relevant for the experiment. During the experiment, the instructions will also be told before use of the program, to make the command use extra clear.

One participant indicated that he would have liked an element of competition, either by performing the exercises together or having a leaderboard. This is not the purpose of the research. Yet, the leaderboard is an interesting future idea. Right now, I do not wish to deceive the participants with fake leaderboard entries; a leaderboard for exercising is most exciting if you would compete against people that you know or look up to and this is not possible as of now, as I do not ask for this personal information during my experiment.

Besides these things, it was asked what the goal of the exercise program is: “Is it to achieve a set number of badges or points?” Sporting has various goals and not a clear one can be set to be the ultimate goal. I wish for every participant to get from the program what he or she is looking for and not determine it for them beforehand.

Lastly, the participants rated the program by filling in Questionnaire 3. The programme was rated positively, as it scored a 3.3 and 3.75 on the GameFlow criteria.

4.2 Final study

First, it was checked whether all t-test assumptions were met before statistical tests were run. This turned out to be true for all variables. Secondly, since three independent samples t-tests were performed on the same dataset, the significant cut-off was adjusted in order to correct for the multiple comparison problem. The Bonferroni correction set this significant cut off at 0.05/3 = 0.0167. Subsequently, the three independent samples t-tests were performed.

The independent samples t-test showed that there was a significant effect for the group condition (GEP or BEP), $t(38) = 3.421, p \leq .001$, with the GEP group ($M = 05 : 45, SD = 01 : 55$) exercising for a longer time than the BEP group ($M = 03 : 58, SD = 01 : 18$). The independent samples t-test on difference in mood showed that there are no differences between the GEP and BEP group in regards to this variable: $t(38) = 0.787, p > 0.0167$. However, the mean mood difference of the GEP group is slightly higher, with $M = 1.550, SD = 1.893$, compared to $M = 0.755, SD = 1.259$ of the BEP group. An independent samples t-test showed that there was a significant effect for the group condition (GEP or BEP), $t(38) = -3.611, p <= .0005$, with the GEP group ($M = 3.4969, SD = 0.29488$) rating the game flow higher than the BEP group ($M = 3.1125, SD = 0.37368$). A Pearson correlation coefficient was computed to assess the relationship between the difference in mood, time of exercising and game flow score. There was no correlation between any of the variables. When the same correlation coefficient was computed, for both the GEP and BEP group seperately, no correlation between any of the variables was found either.
5 Discussion

The first research question was as follows: can differences be found between the baseline and game group in relation to joy, score on game flow of the program and time of exercising? I hypothesised that significant results could be found for all the three variables. Score on game flow and time of exercising differed enough to conclude that participants in the GEP group scored higher on these variables than the BEP group. They experience more game flow and exercise for a longer amount of time, which suggests they are more motivated than the BEP group. No significant results could be found for the score on joy (mood difference). The mean values showed that the GEP group’s mood changed more positively than the BEP group’s, but the difference was not big enough or the group size was too small. A subquestion of the first question was whether differences could be found when the participants would be divided in personal motivation groups. These were statistical tests that I was unable to perform, as only five from all forty participants were not intrinsically motivated. A data split was not possible, as the group distribution would be too skewed to possibly obtain significant results.

The second research question was whether a correlation between duration of the workout, game flow and joy existed. While I thought there ought to be a correlation, none was found. A larger group size might have portrayed different results.

5.1 Limitations and improvements

BEP feedback The best rated feature was the example video, after which voice commands followed and lastly the starting explanation. These features were also the only ones to choose from in the BEP condition. While the example video was best rated, participants who played the BEP most frequently stated that they did not understand the exercise and that the videos needed to be improved. Some videos in the program did not move since no good videos could be found for these exercises: these still videos caused the most confusion. Participants also wished to see each movement largely scaled, before starting a new exercise. Some stated they wished the program would do voice-overs when text appeared, that they could determine how long they wanted to exercise or that they wanted more exercise recommendations from the program. As missing features, almost all participants stated they wished feedback was given by the program, whether it was in text, voice-over, score, 3d stars or how many exercises were left. Feedback and control seemed to be the key in keeping them motivated.

GEP feedback Best rated features for the GEP version were score tracking, sounds, and a tie between the extra points mechanism and achievement badges. From these terms alone, it is clear that the GEP version has a lot more going on to motivate users. It is partly why some participants wished they had more prior knowledge on what they could expect to see on the interface so it would not overwhelm them. Others wanted to know what achievements exist beforehand so they could work towards them. One person stated he would have liked
more challenging achievements, such as earning 50 points within 30 seconds for a bonus. Others wanted more points for difficult exercises, such as Burpees. Now they felt like skipping it quicker because it was a harder way to earn points compared to Jumping Jacks. The skeleton registration trouble of the Kinect also came forward: sometimes a coin would not be collected and participants did not understand why. This is not an unknown problem: Anderson’s YouMove research also marked this as a limitation [1]. The GEP group came up with features that the program could build upon, such as choosing your own exercises (autonomy feature), competing against somebody (relation feature), more options for different achievements (i.e. completing a certain number of movements for one exercise, a competence feature), a progress bar for the full program and a body check to see if the exercise is performed correctly. These are all realistic features that would be of value if the program were to be extended.

It is interesting to see that the named features to improve include features that meaningful gamification builds upon. While I was unable to implement these features in the time I had, it would be interesting to extend the GEP version for future research and look for differences in a BLAP version compared to a BLAP + meaningful gamification version. I also see potential in creating versions of the program that each focus on a different sports domain, such as dance, yoga or fighting sports. In conclusion, the results found within this research are promising to further explore the potential of Augmented Reality exertion interfaces in combination with sports. There are still a lot of possibilities which have not been explored yet.
References

A Selecting participants - decision tree

![Decision Tree](image)

**Fig. 8:** This figure depicts how participants were screened. Convenience sampling was used to find participants.

B Questionnaire 1 - before experiment

(Based on the Sports Motivation Scale developed by Pelletier [20]. Only questions 1, 5 and 22 have been rephrased to fit the research better, their meaning remains the same.)

First answer the following questions for general information. There are no wrong answers considering your preference to performing sports, so please answer in honesty:

1. What is your age?
2. What is your sex?
3. How many times a week do you practice sports? (0-1, 2-3, 4-5, 5+)
4. What kind of sport do you focus on? (Cardio, Strength)
5. Do you use technology during exercise? If so, what kind of technology? (i.e. smartwatch, health apps, YouTube videos or other)

Now please fill in the table below and indicate to what extent each of the following items corresponds to why you practise sport.
Table 1: Assessing sports motivation

<table>
<thead>
<tr>
<th></th>
<th>Does not correspond</th>
<th>Corresponds a little</th>
<th>Corresponds moderately</th>
<th>Corresponds a lot</th>
<th>Corresponds exactly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>For the excitement I get from the experience.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>For the pleasure it gives me to know more about the sport that I practice.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>I used to have good reasons for doing sport, but now I am asking myself if I should continue doing it.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>For the pleasure of discovering new training techniques. I don’t know anymore; I have the impression of being incapable of succeeding in sport.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Because it allows me to be well regarded by people that I know.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Because, in my opinion, it is one of the best ways to meet people. Because I feel a lot of personal satisfaction while mastering certain difficult training techniques.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Because it is absolutely necessary to do sports if one wants to be in shape.

For the prestige of being an athlete. Because it is one of the best ways I have chosen to develop other aspects of myself.

For the pleasure I feel while improving some of my weak points. For the excitement I feel when I am really involved in the activity. Because I must do sports to feel good about myself.

For the satisfaction I experience while I am perfecting my abilities. Because people around me think it is important to be in shape. Because it is a good way to learn lots of things which could be useful to me in other areas of my life.

For the intense emotions I feel doing a sport that I like.
It is not clear to me anymore; I don’t really think my place is in sport. For the pleasure that I feel while executing certain difficult movements. Because I would feel bad if I was not taking time to do it. To show others how good I am at sport. For the pleasure that I feel while learning training techniques that I have never tried before. Because it is one of the best ways to maintain good relationships with my friends. Because I like the feeling of being totally immersed in the activity. Because I must do sports regularly. For the pleasure of discovering new performance strategies. I often ask myself; I can’t seem to achieve the goals that I set for myself.
C Questionnaire 2 - before and after experiment

(This is the Sports Emotion Questionnaire developed by Jones et al. [10]. It is used to assess emotion before exercising and afterwards. For the measurement afterwards, participants have to indicate how they felt during the exercise. After data collection, a score on difference in mood is calculated.)

Below you will find a list of words that describe a range of feelings that you may experience. Please read each one carefully and indicate on the scale next to each item how you are feeling. There are no right or wrong answers. Don’t think too long: choose the answer which best describes your feelings.

Table 2: Assessing emotion (before and after experiment)

<table>
<thead>
<tr>
<th>Not at all</th>
<th>A little</th>
<th>Moderately</th>
<th>Quite a bit</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Uneasy</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. Upset</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. Exhilarated</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. Irritated</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. Pleased</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. Tense</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. Sad</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. Excited</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9. Furious</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10. Joyful</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11. Nervous</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12. Unhappy</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13. Enthusiastic</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14. Annoyed</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15. Cheerful</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16. Apprehensive</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17. Disappointed</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18. Angry</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19. Energetic</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20. Happy</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>21. Anxious</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>22. Dejected</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
D  Questionnaire 3 - after experiment

(This questionnaire is based on the eight GameFlow elements as established by Sweetser and Wyeth [25]. Two questions are asked per element.) Below you will find a list of questions assessing game criteria. Do not spend too much time on any one item, but choose the answer which best describes your experience. A few open questions are also asked to assess what features did or did not speak to you the most.

- Rank the following features (on top = most favoured, on bottom = least favoured):
  BEP options: example video of exercise, voice commands, starting explanation.
  GEP options: example video of exercise, voice commands, starting explanation, score tracking, sounds, collection of objects, achievement badges, extra points mechanism.
- Which features could be improved?
- What features do you think were missing?
- Would you use this program again?
Table 3: Assessing GameFlow criteria

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Undecided</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The program grabbed by attention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>I got distracted by all the different stimuli.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>I felt challenged by the program.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>The program matched my own skill level.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>It was easy to understand what I needed to do.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>I feel as if my skills have improved.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>I had control over the program.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>My actions had an impact on the program.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>I felt restricted by the program.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>It was unclear what the goal was.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>I received feedback on my performance.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>The feedback I received was sufficient.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>I was emotionally involved in the program.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>I worried less about everyday life and/or self.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>The program should support multiple players.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>I missed competition in the program.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
E Scoring system

E.1 Questionnaire 1: assessing sports motivation

KEY [20]
- (2, 4, 23, 27)/4 Intrinsic motivation - to know
- (8, 12, 15, 20)/4 Intrinsic motivation - to accomplish
- (1, 13, 18, 25)/4 Intrinsic motivation - to experience stimulation
- (7, 11, 17, 24)/4 Extrinsic motivation - identified
- (9, 14, 21, 26)/4 Extrinsic motivation - introjected
- (6, 10, 16, 22)/4 Extrinsic motivation - external regulation
- (3, 5, 9, 28)/4 Amotivation

Final Score: category of max(mean(intrinsic motivation score), mean(extrinsic motivation score), mean(amotivation score)).

E.2 Questionnaire 2: assessing emotion

KEY [10]
- (1, 6, 11, 16, 21)/5 Anxiety
- (2, 7, 12, 17, 22)/5 Dejection
- (3, 8, 13, 19)/4 Excitement
- (4, 9, 14, 18)/4 Anger
- (5, 10, 15, 20)/4 Happiness

Final Score: sum(excitement + happiness - anxiety - dejection - anger).

Mood difference score: moodafter_score - moodbefore_score.
A positive difference means an increase in mood, a negative difference means otherwise.

E.3 Questionnaire 3: assessing GameFlow

KEY
Questions 2, 9 and 16 are asked in an inverse manner - these questions are converted when calculating the final score to give the correct values. The final score is equal to the amount of given points summed together, then divided by the total amount of questions (i.e. 65/16 = 4.0625). It represents on a scale of 1 to 5 how game flow was felt by the participant. The maximum amount of points to give is 80 (i.e. 80/16 = 5).