

# Testing Methods for Mobile Game Development

A case study on user feedback in different development phases

Elina Annanperä, Jani Yli-Kantola, Tanja Sauvola,  
Kari Liukkunen  
M3S Research Unit  
University of Oulu  
Oulu, Finland  
firstname.lastname@oulu.fi

Samuli Heinonen, Erkki Siira  
Data-driven Solutions  
VTT Technical Research Centre of Finland Ltd  
Oulu, Finland  
firstname.lastname@vtt.fi

**Abstract**—Mobile games are relatively new industry, and their relevance on the game markets are rising. At the same time, it is very competitive market, as new games are coming out by hundreds every day. The challenges for the new game development come from the creative process and the human factors on how to make the game appealing. In this research, we study the use of test groups as part of new mobile game and concept development process in the early phases. They were used to complement the testing automation at the development phases when use of analytics was not yet possible. Six separate feedback collection sessions were arranged using different methods. The paper presents the main learnings of each used method, the perceived usefulness of the methods for the entire development process. We also contribute to the growing knowledge for mobile game development, focusing on methods for the testing with test user approaches. We conclude that testing with user groups can help clarify the mobile game design to create a more appealing game, but choosing the right groups and right moments in the game development phase can be difficult, as can be the interpretation of the feedback for decision making.

**Keywords**—mobile games, development process, game testing, users, methods

## I. INTRODUCTION

Mobile games take up currently 42% share of all game industry (total of \$108.9 billion), and the market share is expected to grow even more in the next few years [1]. At the same time, the mobile game markets are incredibly competitive. According to a report by Deloitte [2], more than 500 new mobile games come out each day, which means there is a challenge to get the visibility and downloads to a new game. At the moment, familiar entertainment franchises and big game companies, who have the big marketing budgets, are raking in the big revenues. Developing a new mobile game that finds its niche in the market is not easy, and there are no straightforward guidelines for creating a successful game.

This is partly because developing the mobile game is not just software development but requires plenty of different skills and knowledge from the development teams: sound and art, artificial intelligence, control systems, and especially the understanding of the human factors [3]. This makes the game development process a complex task. For the game development to be effective and productive, the development process should be clearly streamlined for the development

team. Testing is one significant phase of the game development process and perhaps the most expensive one [4]. Even though the simplified view of testing is to identify bugs and help developers to remove them, in practice it is not so straightforward [5]. It is estimated that the testing phase can take over 50% of the development project resources [4]. Effective mobile game testing derives from a well-structured and systematic approach, use of test automation framework and seamless integration with development process. In professional game development process, the aim should be efficient and result-driven testing.

On the other hand, Prystupa-Rządca [6] state that game development is a risky business venture due to rapidly changing industry trends and nuanced customer preferences. When testing the mobile game, the entire game concept, and its appeal are tested and evaluated either directly or indirectly. To be able to recognize what is the actual appeal of the game is vital insight prior to launch. To minimize risks of game development, companies typically test their products prior to official launch. For small companies, the need for a rigorous testing phase is even more essential, as they depend much more on the success of each individual game than large companies do. Often, small companies invite individuals from outside the organization to test their products, using different strategies of their implementation onto the project [6].

In this paper, we study a mobile game that is being developed as part of a larger concept. We further focus on identifying the testing needs of the mobile game in its early phases, when the development is much tied to the entire new concept. The main research problem was to identify the suitable methods and testers groups for each phase, collect feedback of the game in its early development phases, while considering the game developers' needs. The research effort focuses on the human aspects and player point of view of game and concept testing. The study was conducted during a research project, involving two games companies, a company building the physical exercise aspect of the concept and two research organizations. One of the game companies was also the main developer and owner of the entire digitized exercise concept.

Rest of the paper is organized as follows. First, we present a relevant theoretical background on mobile game developing and testing. Second, we present the research context and methods, followed by the study results. Finally, the results are discussed and conclusions presented.

## II. THEORETICAL BACKGROUND

### A. Mobile game development

Mobile game development is sometimes thought to be mainly software development, but it is in fact much more. According to Aleem et al. [3], the multidisciplinary nature of the processes that combine sound, art, control systems, artificial intelligence (AI), and human factors, makes the game development practice different from traditional software development. Because of the multidisciplinary activities and competences required, game development is a very complex task. Thus, creation of games involves cross-functional teams including designers, software developers, musicians, script writers, and many others. In order for the game development work to be effective and productive, the development process needs to be streamlined within the development team. Development teams can get benefits by understanding and following good practices of traditional software engineering, but challenges may appear nevertheless (see e.g. [7];[8];[9];[10]).

Mobile game development is relatively new industry domain. Thus, established development processes in the scientific literature are not that many. If including more general processes, the game development process is indeed discussed in scientific literature. Aleem et al. [3] have made a systematic review about game development software engineering process (GDSE) life cycle. They state that the process is usually split in the following main phases: The pre-production phase includes testing the feasibility of target game scenarios, including requirements engineering marketing strategies. The production phase involves planning, documentation, and game implementation scenarios with sound and graphics. The post-production involves testing, marketing, and game advertising. The similar video game development processes breakdown are introduced in several sources, some of them add game concept development in the beginning of the process (see e.g. [11];[12];[13];[6]).

The list below illustrates all actions, which can be seen to appear in those phases.

- Concept (Game mechanics, Setting, Technology, Interaction)
- Pre-production phase (Management, Requirement specification, Game system description language, Reusability, Game design document, Game Prototyping, Design tools, Risk management)
- Production phase (Asset creation, Storyboard production, Development platforms, Formal language description, Programming, Game engine, Implementation)
- Post-production phase (Quality assurance, Beta testing, Heuristic-based testing, Empirical testing, Testing tools, Marketing)

Aleem et al. [3] bring up that due to relatively young history and empirical nature of the field, there has not been any development strategies or set of best practices to carry out game development fully explored. Their study helps to identify

the research gaps in game development life cycle. They suggest that researchers should pay more attention to the game development process, especially in the phase of post-production.

Overall, as the mobile game development research is still maturing, the less scientific blogs, white papers and other more casual forums discuss mobile game development processes, best practices and tips for game developers. For instance, Dulskiy [14], an experienced game designer states in his blog that an average mobile game development process has distinct phases: from idea and concept development to developing proof-of-concept, creating game design document and prototypes, to designing the architecture, game development, testing the game and finally support.

All the descriptions discussed here about the game development process can be seen as simplification of the process and they look like they are following a traditional software development model. In order to understand the mobile game development, we combined features of the processes discussed in these different sources and outlined them to a development process, shown in figure 1. The processes described here seem to suggest that generalizations of design, development and testing phases can be made. In our depiction, we wanted to emphasize the iterative development, where game design starting from the idea and concept design advance to development (prototypes and finally architecture) to testing the mobile game in different ways in different stages. The iterative approach was chosen, because it seems that the overall game development is not suited for typical software life cycle methods, such as the waterfall model. This view is shared by O'Hagan et al. [15] in their literature review on software process models used for game development. They state that agile and hybrid approaches are used by most organizations for game development. Agile approach is based on iterative prototyping, a subset of software prototyping. Agile

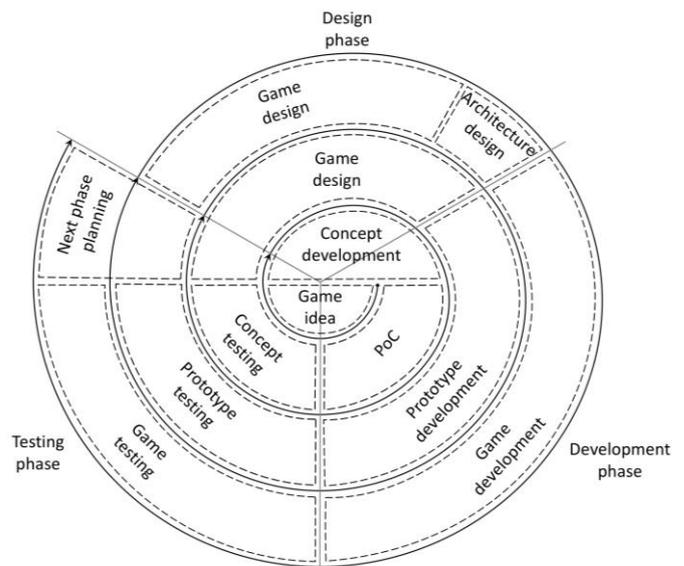


Fig. 1. Mobile game development process (authors' own depiction)

development depends on feedback and refinement of game's iterations with gradually increasing feature set (see also [16]).

### *B. Mobile game testing*

According to Kasurinen et al. [4], testing is perhaps the most expensive task of any software project. They present an estimate that the testing phase took over 50% of project resources. They also point out that besides causing immediate costs, testing is also importantly related to costs related to poor quality, as malfunctioning programs and errors cause large additional expenses to software producers. Kasurinen et al. say that the testing work can be divided into automatic testing and manual testing. Automation is usually applied to running repetitive tasks where test cases are executed every time changes are made to the product. Manual testing suits better in tasks where is only little repetition, such as explorative testing or late development verification tests [4]. Test automation requires the utilization of testing tools or writing test scripts, so the actual testing work is done by computers. Manual testing is done by humans, which means in this case end-user testing.

Mobile game development being such a new industrial area, Aleem et al. [3] stated that post-production phase still have research gaps, when it comes to the development life cycle of mobile games. Post-production phase includes activities like quality assurance, beta testing, empirical testing and utilizing testing tools for test automation.

Helppi [5] in his blog for BitBar (a mobile test automation company), states that mobile game testing differs from the regular mobile app testing. According to him, effective mobile game testing derives from a well-structured and systematic approach, use of test automation framework and seamless integration with your agile process. Because mobile game development is not just software development, also the testing of mobile games is different compared to the traditional software artifacts. Murphy-Hill et al. state that game developers are hesitant to use automated testing because of these tests' rapid obsolescence in the face of shifting creative desires of game designers [17]. Moreover, Kasurinen and Smolander suggest that in the testing phase game developers should focus on soft values such as game content or user experience, instead of more traditional objectives such as reliability or efficiency [18].

The simplified view of testing is to identify bugs and help developers to remove them. In professional game development process, the aim should be efficient and result-driven testing. Thus, it is mandatory to have test automation as a part of the overall development process. The following six key areas are listed in mobile game testing which should be tested carefully: User Interface and functionality, Graphics performance, Usability and User Experience, Multi-player / User features, Social integrations, Security and Liabilities.

Helppi [5] also list the most important testing methods to be used in mobile game testing. Functional testing is associated with manual testing and playing 'game through'. Automated functional testing can reveal issues related to user interface, stability, game flow and mechanism, and integration of graphics assets. Compatibility testing is all about making that game compatible across different devices. Performance testing

ensure that game runs at least in the most of the mobile devices in the markets, not only in the high-end devices. Localization testing is important if targeting to the global launch, meaning that all texts, titles and content is localized and screen layouts fit perfectly to the variable content. Regression testing needs to happen when anything changes in the software. Especially the server-client interaction must be ensured that either server side or client side code changes do not break the functionality of the whole service. Load testing covers tests which measures the limits of the system, such as the number of players on a server, the graphic content on the screen (e.g. frames per second, FPS), or memory consumption of the system. Helppi put weight to test automation and highlight that many of these testing methods should be done or supported by automated testing.

Helenius et al. [19] interviewed in their study Finnish game development companies to gain understanding of the success factors of developing flourishing games. They defined patterns, which are certain models of behavior in game development process. The patterns are meant to help reaching goals in development projects. Relating to mobile game testing, they emphasize the user perspectives and investing resources on user centric evaluation of the game. They mention such practices from gaming industry such as heuristic evaluation (following Nielsen [20]), using outside experts (peer evaluation) and shadowing. The latter means that the game is given to test users to play around with, for the need to have an insight on how the testers react to the game. Helenius et al [19]. point out that the user needs to feel comfortable with the game within the first 15 seconds of starting with it.

On the other hand, one of the most common way to test a mobile game is establishing an alpha and beta testing groups from active mobile gamers and ask them to test an early stage game releases. While gamers are testing the game, the game analytic data is gathered from the most important game events. Helenius et al [19] reported that based on their observations game companies value the most retention rate. The reason is that the more players come back the more they like the game. According to their study, the other important statistics are session times and game events as these reveal problematic designs inside the game. McCalmont [21] lists the most important mobile game analytic metrics. They are e.g. Daily Active Users, Sessions, Retention, all of this are related to the game functionality and mechanics and are seen as the basic data metrics. Furthermore, there exists metrics, which are more related to the game monetization. These metrics are Conversion Rate, the Average Revenue Per Daily Active User (ARPDau) and Average Revenue Per Paying User (ARPPU). All of these metrics are gathered during the testing phase. Based on this data set, game developers can make certain analysis of the game operations and possibly make some forecasts how the game in question could succeed in the markets. This is quite solid and proven testing method for mobile games when they are coming near the soft launch or actual global release.

Even though game analytics is quite powerful method, it has some shortcomings. The most important deficiency is that it does not necessarily tell the reason behind some outcomes. For example, it is easy to see from the data if players have stopped their game sessions at a certain point or after specific

amount of time, but it does not tell whether the reason is in the game graphics, game mechanics or in some other game attribute. It should be noted that in order that game analytics can be used, there must be an executable game version available for testers. Thus, in this study we expanded the game testing by user centric approaches to testing process to get deeper understanding for gamers’ opinions about the game for several different release versions. User centric testing also gave us a possibility to ask development ideas for the mobile game version in question.

### III. METHODOLOGY

#### A. Mobile game and new exercise concept

The context of this study is a new business concept that aims to combine mobile gaming with real life exercise. The concept includes the mobile game that links real life exercise and digitalization of exercise. The mobile game is designed to include data from exercise through activity sensors and have a location-based map feature. The game also includes an exercise themed mobile game that keep the players coming back to the game, use all the offered features and get the players slowly exercising without the feeling of being pushed to do so. The game guides the player to exchange real life activity to the game energy in the mobile game, and on the other hand gain rewards from the game that can be used in real life parts of the concept.

The project consortium included two game companies: a startup that owned the exercise game concept idea and was developing the mobile game, an established larger game company that acted in a mentor role in the context of this research. Additionally, there was a company focusing on creating the physical exercise part that links to the game in this concept, and two research organizations with expertise in both the technical aspects and also the user centric aspects of the exercise concept and mobile game design and development.

In the beginning of this study, the concept idea and rewarding logic existed, but the game design was not finalized. At the start of the testing activities, we had a version of the game that had some of the game elements ready, such as layout graphics of different spaces, basic game features and basic map feature for real life exercise. The player could choose a character of their liking from different options, and change of

its features, such as clothes. The focus of our study was to support the game design and concept development with user centric approaches. As the game development was in its early stages, analytics data was not available until the last phases of testing. Thus, in this study we expanded the game testing by user centric approaches to get deeper understanding for gamers’ opinions about the game for several different release versions. User centric testing also gave us a possibility to ask development ideas for the mobile game version in question, and for the overall concept.

#### B. Research process

The game development and testing process in this research is presented as a case study [22]. The concept under development was seen as a novel case for mobile game development, as it was closely tied to the overall exercise concept. The study follows the game development process for a year’s time span, from the first playable game version to soft launch in Finland. The feedback collection was planned together with the game developers to help in each stage of the game development process. Altogether, we arranged six separate data collection sessions. In table 1, the events are summarized.

First, in the **expo release**, there was need for introducing the game and the concept to potential audience and collect general feedback on one hand from the general concept and on the other hand from the game itself, the first impressions. This event was organized at a fitness expo event over two days, and participants were recruited on the spot. The testers were expo visitors, interested in fitness and bodybuilding, mostly young adults. Each tester was given the game to test (on a tablet device), given some basic tasks to perform in the game and the reactions and reflections were observed and written down. Then, the exercise concept was presented and discussed with the tester. Finally, they were asked to fill in a survey that had some detailed questions on the game and concept. One test session lasted about 15 minutes.

The second event, **focus group** (following [23]), was two months later, taking place at the university part of this research. The aim was to test the renewed start of the game and get first time reactions from it. Additionally, we wanted to test the use of focus group approach to collect feedback. The participants were university employees of ICT-related background and

TABLE I. SUMMARY OF DATA COLLECTION

#	Release	Distribution	Aim	Method(s)	Participants
1	Expo	Internal	Feedback of concept, first impressions of the game	Test scenario, Observation, survey form	expo visitors (62)
2	Focus Group	Internal	Feedback from “first 20 minutes” gameplay	Focus group: observation, discussion	University staff (4)
3	Expert Workshop1	Internal	Feedback of new features, Test the feedback collection through the game	Observation, discussion	Project consortium (facilitators 3, testers 7)
4	Expert Workshop2	Google Play Alpha	Feedback of finalized features, Test the feedback collection through the game	Observation, discussion	Project consortium (facilitators 3, testers 7)
5	Beta	Google Play Beta	Feedback from first impressions of beta-version	Focus group: observation, discussion	Students (11)
6	Soft launch	Google Play	Feedback of experiences of longer gameplay	Group interviews, survey form	Students and technical researchers (30)

were not familiar with the game or the concept beforehand. The testers were asked to play, and the situation was observed, and notes taken. When the testers had reached a certain point, they were asked to stop. Then the experience was discussed and notes were taken. There was one main facilitator for the test, but each tester had a corresponding observer. The session lasted two hours.

Two consortium workshops, **expert workshops 1 and 2** in table 1, with similar setup took place within four weeks of each other. They were needed to internally test the new directions the game had taken to collect expert feedback from people already familiar with the concept and previous game versions. In these workshops, the participants were given the new game version to test and observation method was used to collect the reactions and comments. In both workshops, the game testing lasted about 45 minutes. After the game session, the contents, ideas and other feedback was once more discussed.

**Beta release** testing was arranged two months after the latter internal workshop. Two polytechnic gaming student groups were recruited and they were not told about the concept or the game before hand, to get their true first reactions. The participants were asked to follow the game tutorial and then find and try certain aspects of the game. After about 45 minutes, the feedback discussion was held. During the feedback discussion, also the entire exercise concept was presented and discussed.

National **soft launch** release was used two month later to collect new data. Six groups of various types of participants were recruited to first play the game for a week, fill in a survey, and then come to a group session to be interviewed. These groups consisted of fellow researcher colleagues that were not familiar with the game, student groups of various background: information systems students, information processing students (two different universities and one polytechnic) and nursing students (vocational education). These group interviews lasted 1h to 1h 30min each, and were thematically organized to cover the aspects of the game, but also the collect information about the exercising connection related to the game, as well as the players' feelings for game: what aspects were the most fun, frustrating, confusing or addictive. The survey was used to collect background information and feedback of the basic features in various scale questions. We looked into the Game Experience Questionnaire (GEQ) tool [24] for game evaluation, but decided against using it directly, because we felt it was quite heavy for the testers to answer in our study design set up (i.e. group interview) and it would have to have been translated into Finnish. It was mostly used as inspirational source for some questions on testers feelings towards the game.

These research cycles, their aims and outcomes were discussed and verified with the game developer company. This was done in order to ascertain the phases and their outcome from the company point of view, and get the developers' opinions on the benefits of each phase. The results and learnings of the testing cycles are presented and discussed next.

#### IV. RESULTS

The identified phases in mobile game testing with users started with the release and distribution channel decisions, and

followed by choosing the methods used in each phase, analyzing the outcome and finally decision making on the next design iteration. Some of the needed decision were made inside the game development company (how to distribute, planning the next phase), some in collaboration with the researchers (release schedule) and some of them, the researchers made the decision (choosing methods and the analysis). The results of the phases are elaborated in the following.

##### A. *Early concept evaluation*

Initially, active, young, fitness oriented people were seen as one of the potential target groups for the concept and the game. Feedback collection based on the **expo release** at the expo event, gave results that were somewhat contradicting to the expectations. The results indicated that the concept sounded and seemed new and different (70% of the respondents in the survey) and therefore interesting. On the other hand, many of the participants (53% of the respondents in the survey) stated that they play mobile games rarely or not at all. Therefore, there is the possibility that an interesting concept will not reach its target audience via the mobile game. The game itself seemed complicated to many testers, and they did not seem very comfortable testing it at the expo all of a sudden. The feedback of the game was that there should be more to do in the game: more content in the map feature connecting to real life events, community building aspects and competing with others.

The concept did received a very positive feedback. The testers thought that it could encourage gamer-oriented people to exercise. The expo audience mostly were interested in the rewarding system, and saw the game a potential platform for converting their already active lifestyle to rewards. Specifically they mostly wanted to get rewards in the form that would support their exercise hobbies.

After the feedback and experiences were shared to the company, they felt the feedback was encouraging in the sense that the testers were interested of the concept, though maybe found the game not meeting their expectations. It was decided that the next phase of game development should focus on the introduction to the game, in order to lessen the confusing feeling when starting the game in this early phase. The starting of the game was seen important for creating the feeling for the game, motivating the player to invest time in playing and exploring the available features.

Next, the **focus group** release feedback was collected. From observing the testers, it was found that even with only four people, there was variation in the speed testers progressed, and how much time they spend going through the introduction. When discussing about the experience, the feedback was mainly positive, the game looked good, and the idea was the most positive aspect. On the other hand, the gaming experience was not very enthusiastic, the testers were not sure if they would play the game on their own, and the execution of the idea seemed complicated at this release. The testers expected that the game would have a stronger link to real life exercising. Because, there was not such indication at the game introduction at this phase, they were somewhat critical about the purpose of the game.

There were initially plans to arrange more of these types of groups. However, when going through the feedback with the game developers, it was decided that the game mechanics needed more focus. The developers also felt that the user feedback was not helping them much to develop the game in this respect. This was because the testers could not really give insight to improve the game, they can only tell what they like or do not like, and give suggestions based on the things they see. It was decided to focus on the project internal, expert feedback to improve the game.

### *B. Gameplay evaluation*

In the **first expert workshop** within the research project consortium, the changes made to the game mechanics were seen as the step towards right direction. The main evaluation from the expert testers was, that on the positive side, there were more competitive elements and the game graphics were mostly pleasing, and the general impression was that gameplay aspects showed promise. Still the introduction to the game story took too long. Feedback collection via a form that opens from the game was tested during workshop. Method was not seen user friendly to be used for wider distribution, but could be used on small, focused testing groups.

At this point, the game developers decided they needed to focus on the things that work in the game and finalize them. These were the minigames and the tournaments where the game character competes against others to gain experience and resources. The main game features were present at the game, but some of the game mechanics needed to be tested still.

At the **second expert workshop**, the researchers tested another variation of collecting feedback directly while playing the game. We used embedded questions that come at game transition phases, trying not to disturb the immersion too much. The game's playability aspects had been improved and new game mechanics introduced. As result of the workshop, it was agreed that the game could soon be beta launched and tested with different groups of people.

Preparing for the release, the game went through some changes. The graphics were updated, game included two different minigames and tournament logic. One game avatar, a male character was included, and there were possibilities to modify the character's clothing and his home. The changes had made the game simpler, though there were still the exercise related features: map and sensor tracking. The tutorial of the game guided the player through all the features: first play the minigame, then the tournament, and how to utilize exercise features for replenishing the game energy. From the research perspective, it was decided, that the feedback collection through the game would not be used at all. It was seen as a distraction, and the game developers and the researchers were both worried it would alienate people from the game.

### *C. Holistic evaluation*

The **beta testers** seemed to get into the game quite fast. The minigames got good feedback: the fast paced and easily playable games were fun in the testers' opinion, and they also liked the competitive aspect. On the other hand, they stated that there were many different features to go through in the game

tutorial, and it mostly just confused them. They would rather just explore the game more by themselves. From the concept perspective, a lot of the feedback focused on how the exercise aspects of the game could be more prominent, and what other features the map and the real life exercise link could have. The main benefit from the feedback and ideas was that it gave emphasis and validation for the next steps.

For the national **soft launch**, some additional content was added, such as a new minigame, and new sensors to link to. There was now a shorter tutorial, that would not showcase all the features at once, but left something to be discovered for the player and kept the idea of the game simple. The game developers wanted more insight of the general interest of the game, if players thought it was actually funny and "addictive", and what were the most popular or liked features. They were also interested in the concept expansion ideas and developing the rewarding system. It was agreed that this test round would include people that had already played the game for a while.

As could be expected, there was variation on the time the testers had actually spend playing. The game and concept received a mixed reaction. Others had found the minigames in particular fun and played to the higher game levels quickly. Others had stopped after or during the tutorial. The game concept received again good feedback: it was considered new and interesting and to have staying power. The game looked appealing. The most liked aspects were the minigames: they were easy, fast paced, and gave a sense of achievement. Surprisingly, only few of the testers had spent really time on the exercise features or connected their fitness sensors to the game. Rest of the testers told that they had not really noticed these features, or when tested them had been indoors and not able to test them properly. Most of them did suggest that if increasing one's activity was the point of the game and concept, these features should be more visible. When presenting the upcoming rewarding mechanism to the testers, some suggested that the game could be used personal goals, and rewards based on achieving the goals.

The developers also had at this point some analytics data they followed. Combining the feedback of both sources, the changes for extended soft launch, (expanding the launch to new countries) were decided. The game character would be changed, and both female and male character choice given. The graphics were to be completely updated (from 3D graphics to 2D, for faster graphics production), and new minigames would be added. A bigger new feature was ready to be included at this point: new rewarding system via collaboration with third party vendors that would show as an interest points in the map feature and give players access to special promotions in exchange of the activity data they collect in the game.

## V. DISCUSSION

Mobile game development is not an easy and straightforward endeavor. It is even more challenging when there is a predefined bigger picture, such as in this study, the new type exercise concept, where the game needs to fit in. In this study, we have followed the development of one such mobile game by arranging testing sessions that correspond to the development phase of the game and the concept. Early on

in the game development, there was the goal of combining real-life exercise and gaming. How to best achieve this goal, while making the game interesting and fun, was the challenge in the development.

The first two tested releases focused on showcasing the concept, building the storyline, and introducing the game characters, assuming that those would be the most interesting aspects of the game. The game mode resembled those of strategy games. Overall, the testers were very interested in the concept idea. However, as a game, the testers wanted to have more active playing, and a stronger link to their own exercise. Discussions on how to turn the feedback into development plan were had with the game developers. In a sense, the entire game idea and concept was looking for direction in the beginning, and the outsider comments were interpreted so that transferring from story driven game to mini-games driven game was needed, while developing the real life exercise aspects.

The project consortium wanted to support the game development, and as they had the shared view of the concept, it was deemed most purposeful to continue testing the game internally for a while. This also allowed for shorter-term preparation for test sessions, and direct discussions on the direction of the development. The developers were testing a lot of ideas, and needed the tester to understand that they were not looking at a final product. When the confirmation for the maturity of the game was gotten, it was polished for beta testing and for soft launch. The whole process from first testable version to soft launch took almost a year, and this case shows the challenges of combining the elements of a new game and concept together in an interesting way.

The main challenges in getting useful feedback are related to the game development stage. In the beginning, the concept and the game both needed feedback. Separating the two more distinctly was not seen important for the developers as they asked the researchers for help to get the regular players' first reactions and feeling for both (game and concept) in those game releases. Collecting this type of feedback was challenging in terms of further analysis. Separating the game and the concept more distinctly, however, would have in retrospect helped in the situation in the phases, when the feedback was difficult interpret. Testing even smaller parts of the game could have helped in figuring out the features and logic more effectively. In this sense, though test planning was made early on the development, we should have made it more precise, as is suggested in the literature (see section 2.2) and tie it more closely to game design.

However, it is worth noting that in any case, interpreting the tester feedback can be difficult: when something does not seem right or interesting enough for the tester, they may not be able to tell exactly why it is so, or how to fix it. This is partly because the testers were focusing on things they see in the game are not necessarily very good at telling why they feel the way they do. This is the case whether they like the tested aspects or not. It was up to the interpretation of both the researchers and the developers to try to get to the actual reasons behind the feedback, and base the decision-making on that. Testing the same feature with different variations would be a

suggested solution to help focusing the feedback. However, testing takes resources that small company making their first game does not necessarily have, and they have to make the choice on where to put those resources.

This early on the development, the methods for collecting feedback were limited to interaction directly with the testers, thus not being able to get that big of a data set to benefit from analytics data. Data collection methods such as video recording the test sessions, when outsiders were asked to test the game was considered. In the case of the expo visitors, it was considered too difficult to do, as the expo hall was quite noisy. In the focus group release, if those groups would have been continued, it would have been a good way for collecting data. Later, the survey questions were tested among the expert group, where the approach did work, but was feared to be too distracting to regular players. Observation was considered the most effective ways to get the first reactions and overall impressions of the testers in both the expert workshops and in the beta testing groups. Recording conversations and using a separate survey to add to the interviews were easy to set up and later analyze, and did not seem to bother the testers at all.

As has been stated previously, the concept was very well received by testers in all the different phases. The tester feedback on one hand reassured us that the development was going in the right direction, but also gave some new ideas where the concept could be extended in the future. It seemed that targeting the game to already active people, that have interest in gaming they benefit from the rewarding system. Variations of the game can be released to target different user groups. It remains to be tested exactly how easily the player is nudged into exercise if he is not prone to it already. But using the game a platform to organize campaigns with the 3<sup>rd</sup> party organizations offering promotions through it, we could study the actual health benefits. In these instances, that testing methods need to be again thought to fit the use case and audience.

## VI. CONCLUSION

In this study, we conducted mobile game testing as part of a new mobile game development process. The additional challenges to the mobile game development were that the game was designed to be part of a larger exercise concept combining real life and digital exercise. The game development was in the early phase, and testing focused in the early releases almost as much on the overall concept as on the game. However, utilizing user centric feedback collection methods, arranging feedback session using different data collection methods and test groups, we were able to collect feedback, ideas and pinpoint challenges for the game developers. The developers were able to make game design decisions based on the feedback, and different feedback collection methods gave us valuable experience in how to choose the appropriate method.

Overall, when developing an entire new concept, the idea needs to come across in a simple way in a mobile game. Although the developers felt that the game was showed to the larger audience too early for them to get valuable feedback, they did realize that the concept was not coming across easily. Testing alternative versions and fine-tuning them was more

effective, when done using expert feedback. Arranging focused sessions with testers in the beta and soft launch release phases, allowed the developers to get some additional insight that the analytics tool was not giving them.

In this research, we were able to build a mobile game development test process for game that is part of a bigger concept. This work is expected to contribute to the maturing research field of mobile game development, but further research on how to effectively test mobile games using user centric testing methods is needed to validate and further explore our findings.

#### ACKNOWLEDGMENT

This work was supported by TEKES (Finnish Funding Agency for Technology and Innovation) and the HILLA research program.

#### REFERENCES

- [1] Newzoo “The Global Games Market Will Reach \$108.9 Billion in 2017 With Mobile Taking 42%.”, 2017 retrieved 22.11.2017, <https://newzoo.com/insights/articles/the-global-games-market-will-reach-108-9-billion-in-2017-with-mobile-taking-42/>
- [2] Deloitte “Mobile games: leading, but less lucrative” 2016, retrieved 22.11.2017, URL: <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Technology-Media-Telecommunications/gx-tmt-prediction-mobile-games.pdf>
- [3] S. Aleem, L.F. Capretz and F. Ahmed “Game development software engineering process life cycle: a systematic review,” *Journal of Software Engineering Research and Development*, vol 4, article 6, November 2016, DOI: <https://doi.org/10.1186/s40411-016-0032-7>
- [4] J. Kasurinen, O. Taipale and K. Smolander “Software Test Automation in Practice: Empirical Observations” in *Advances in Software Engineering*, vol 2010, article ID 620836, 18 pages, 2010, DOI: <http://dx.doi.org/10.1155/2010/620836>
- [5] V-V. Helppi “Mobile Game Testing – Part #1: The Importance and Difference from App Testing” 2014, retrieved 22.11.2017, URL: <https://bitbar.com/mobile-game-testing-the-importance-and-difference-from-app-testing/>
- [6] K. Prystupa-Rządca and J. Starostka “Customer Involvement in the Game Development Process,” *Journal of Entrepreneurship Management and Innovation*, vol 11(3), pp. 43-66, 2015 DOI 10.7341/20151133
- [7] F. Petrillo, M. Pimenta, F. Trindade and C. Dietrich, “What Went Wrong? A Survey of Problems in Game Development”, *Computers in Entertainment (CIE) - SPECIAL ISSUE: Media Arts and Games CIE Homepage archive*, Volume 7 Issue 1, February 2009, DOI: 10.1145/1486508.1486521
- [8] J. Blow, *Game Development – Harder Than You Think*, Queue - Game Development Queue Homepage archive, Volume 1 Issue 10, February 2004, Pages 28, ACM New York, NY, USA, DOI: 10.1145/971564.971590
- [9] M. Washburn Jr, P. Sathiyarayanan, M. Nagappan, T. Zimmermann, C. Bird, What went right and what went wrong: an analysis of 155 postmortems from game development, *Proceeding ICSE '16 Proceedings* of the 38th International Conference on Software Engineering Companion, Pages 280-289, DOI: 10.1145/2889160.2889253
- [10] F. Petrillo, M. Pimenta, F. Trindade, C. Dietrich, “Houston, we have a problem...: A Survey of Actual Problems in Computer Games Development”, *Proceeding SAC '08 Proceedings of the 2008 ACM symposium on Applied computing* Pages 707-711, DOI: 10.1145/1363686.1363854
- [11] R. Edwards “The Game Production Pipeline. What Goes into Making A Game?” 2006, retrieved 22.11.2017, URL: <http://www.ign.com/articles/2006/03/16/the-game-production-pipeline-concept-to-completion?page=2>
- [12] A. Acharya “Mobile Game Development Process” 2017, retrieved 22.11.2017, URL: [https://www.gamasutra.com/blogs/AnishAcharya/20170406/295304/Mobile\\_Game\\_Development\\_Process.php](https://www.gamasutra.com/blogs/AnishAcharya/20170406/295304/Mobile_Game_Development_Process.php)
- [13] A. Bobeshko “Mobile Game Development Process Explained” 2017, retrieved 22.11.2017, URL: <https://game-ace.com/blog/mobile-game-development-process-explained/>
- [14] I. Dulskiy “9 Stages In Mobile Game Development” 2015, retrieved 22.11.2017, URL: <https://artdriver.com/blog/9-stages-in-mobile-game-development>
- [15] A.O. O’Hagan, G. Coleman and R.V. O’Connor “ Software Development Processes for Games: A Systematic Literature Review,” in *European Conference on Software Process Improvement 2014*, pp. 182-193, June 2014 DOI: [https://doi.org/10.1007/978-3-662-43896-1\\_16](https://doi.org/10.1007/978-3-662-43896-1_16)
- [16] J. McGrath “The Game Development Lifecycle - A theory for the extension of the Agile project methodology”, 2011 retrieved 22.11.2017, URL: <http://blog.dopplerinteractive.com/post/112172271166/the-game-development-lifecycle-a-theory-for-the>
- [17] E. Murphy-Hill, T. Zimmermann and N Nagappan, “Cowboys, Ankle Sprains, and Keepers of Quality: How Is Video Game Development Different from Software Development?” *ICSE'14*, <http://dx.doi.org/10.1145/2568225.2568226>
- [18] J. Kasurinen and K. Smolander, “What Do Game Developers Test in Their Products?”, *ESEM '14 Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, <https://doi.org/10.1145/2652524.2652525>
- [19] M. Helenius, P. Kettunen and L. Frank “Learnings from the Finnish Game Industry” in *Proceedings of the 10th Travelling Conference on Pattern Languages of Programs* Article No. 12, April 2016, DOI: <https://doi.org/10.1145/3022636.3022648>
- [20] J. Nielsen, *Usability Engineering*. San Diego: Academic Press, 115-148, 1994.
- [21] T. McCalmont “15 Metrics All Game Developers Should Know”, retrieved 22.11.2017, URL: <https://gameanalytics.com/blog/metrics-all-game-developers-should-know.html>
- [22] P. Runeson and M. Höst “Guidelines for conducting and reporting case study research in software engineering,” in *Empirical Software Engineering* vol 14 pages 131-164.
- [23] R. Curedale, *Service Design, 250 essential methods*, Los Angeles: DCC press, 2013
- [24] W.A IJsselsteijn, Y.A.W de Kort and K. Poels “The Game Experience Questionnaire. Eindhoven: Technische Universitet Eindhoven, 2013