OULUN YLIOPISTO
UNIVERSITY of OULU

Methodologies for evaluating the playability of mobile games - Systematic Literature Review

Abstract

The gaming industry has been growing rapidly during the past years due to the interest of the new generations in mobile gaming. To deliver a great experience for the gamers, it is required for the gaming companies to produce games that are challenging but at the same time easy to play. To achieve this, it is required to understand the factors that affect the gaming experience. Playability is a term that is used to understand the usability of a game and its experience.

The purpose of this thesis was to understand what is known related to the playability of mobile games and to identify the methodologies that are used by the community to evaluate this phenomenon. To find the answers to these questions, it was performed a systematic literature review (SLR) using the databases Scopus, IEEE Xplore, and Web of Science. After conducting the SLR, 1,390 studies related to the playability of mobile games were found from which 27 were identified as primary studies of this research.

From the data collected from the primary studies, there were identified 12 different methodologies that are used for evaluating the playability of mobile games. The methodologies that are most suitable to assess the playability of mobile games are heuristic evaluation and playtesting. Other methodologies can be used for evaluating the playability of mobile games, but they must include a set of heuristics that allows evaluating the playability. The limitations of the research were mentioned, and it was proposed topics for future research of this field.

The contribution of this thesis is the summarizing of the current methodologies that are used to understand and evaluate the playability of mobile games. The results of this thesis are valuable for game developers, game designers, and game usability practitioners.

Keywords
Playability, game usability, video games, mobile games, systematic literature review

Supervisor
PhD, University lecturer Mikko Rajanen

Foreword

Contents

1.   Introduction

Society is shifting into a more digital world, and that also includes entertainment. The game industry has increased over the years. Thousands of games are created every year and new gaming companies are being established. Furthermore, the number of gamers is growing. According to Newzoo, a games market and insights analytics company, there are around 2.7 billion gamers (Wijman, 2020). Furthermore, the current situation with coronavirus lockdown has caused a growth in interest in gaming, particularly in mobile games (Wijman, 2020).

Games are the most used mobile app (Purcell,2011, as cited in Khan et al., 2015) Although the main purpose of mobile phones is to communicate, nowadays mobile devices are used for all kinds of reasons, including gaming. However, mobile devices are not designed for this purpose, this makes the experience of playing video games on mobile devices awkward and even hard to play (Wong et al., 2010).

Besides free games, mobile users also are looking for games that have high quality and high usability standards. For a game to be successful in the game market, adequate game usability and an attractive game interface are a must. However, it seems that the use of a good usability principle is ignored by the game developers. The purpose of this is to reduce development time and cut corners with the aim of launching the game as soon as possible before the competition. This can have a negative impact on the quality of the game. Thus, game companies would beneficiate from testing thoroughly the usability of the game before publishing it into the app store. (Khan et al., 2015)

For the mobile game industry to provide great games that are fun to play, instead of a painful experience, it requires the game to be entertaining and easy to play. However, it also requires to be challenging enough that the game is interesting to play. Also, it should not be so complicated that the user gets frustrated and does not want to play anymore. (Korhonen, 2016)

To achieve this balance between challenging but easy to play, it requires the game developers and game designers to follow good usability practices that allow the players to go through the game without any hassle (Korhonen, 2016). Korhonen (2016) describes the success of a video game as measured according to the experience it provides to the players. If the players find the game enjoyable, then they will continue to play it. On the other hand, failure of designing games with proper usability issues and bad user experience creates a negative effect on the quality of the game and therefore its success (Pinelle, Wong, & Stach, 2008).

Due to the differences between games comparing to other kinds of software, it is not possible to use the traditional heuristics for evaluating usability in the video game context (Sánchez et al., 2009). For example, the goal of traditional software is to make the user do certain tasks easier. On the other hand, the goal of games is to entertain and challenge the users (Korhonen, 2016). Therefore, a new term was conceived to describe the usability of games, playability. A better understanding of playability allows game developers and game designers to create better experiences for video game players and therefore create more successful games.

The purpose of this thesis is to identify the methodologies that are used by professionals to evaluate the usability of games (playability) for mobile devices and to understand which are the most effective methodologies that are used to evaluate the playability of

mobile games. To be more specific only games that are played using the mobile device. Therefore, games that use virtual reality (VR) devices, wearable devices, or other kinds of gadget besides the smartphone/tablet will not be part of the scope of this study. Since the purpose of the study is to evaluate the playability of mobile games, only games that are able to be played by the user will be included in this research. Thus, primary studies that contain information about not functional prototypes will not be taken into account. To achieve this goal, a Systematic Literature Review (SLR) is conducted for the purpose of getting the most information available related to this topic.

The thesis is structured in the following way: The second section provides background information of this thesis which contains knowledge related to the game industry, usability, and playability. The third section presents information about SLR and how it was conducted for this thesis. The fourth section describes the results obtained from the SLR. The fifth section contains a discussion about the thesis. The sixth section presents the conclusions and as well includes suggestions for future research. The seventh section contains the references used in this thesis.

2. Background and Related work

2.1 Video Games and game industry

During the last years, there has been a booming in the game industry. Thousands of games are created every year and new game companies are emerging. In just 2020, 10,623 games were published in the online gaming platform Steam (Clement, 2021). At the same time, the number of players is increasing at a fast pace. In the year 2020 the number of mobile players is expected to be 2.6 billion, of which 38% will be willing to buy the game (Wijman, 2020). According to Sánchez et al. (2012) currently, the video game industry is the most profitable entertainment industry.

Video games, also known as digital games, are a type of game that consists of some sort of computing machinery, such as computer games, mobile games, or console games (Dörner, Göbel, Effelsberg, & Wiemeyer, 2016). Video games are a type of art that combines traditional elements (such as music, storytelling, and visual arts) with modern technology (Mylly et al., 2019). The main goal of video games is to entertain the users and create different kinds of experiences and emotions (Korhonen, 2016).

The games are composed of three components: a player who wants to participate in the game for amusement, rules that define how the player will interact with the game, and goals that the player needs to achieve to complete the game (Mylly et al., 2019). From a game-centric approach, according to Schell's model, the four elements of a game are mechanics (the rules and goals of the game), story (the sequence of events that occur during the game), aesthetics (the presentation of the game including game graphics and quality of sound), and technology (the platform that is used to deliver the game) (Schell, 2008, as cited in Korhonen, 2016).

2.2 Mobile games

The way we use mobile devices has changed during the last decades. Before, the main purpose of phones was to communicate with others. However, the advances in technologies, it has broadened the things that you can do with phones, including gaming. Montola, Stenros, & Waern (as cited in Korhonen, 2016) defines mobile games as "games that extend gaming to the real-world environment and mobile devices are used to deliver information between the physical and the virtual worlds" (p. 12). One of the main characteristics of mobile games is that they can be played from anywhere at any time (Korhonen, 2016). According to Ponnada & Kannan (2012) the main characteristics that make mobile games entertaining to play are its content, storyboard, rewards, graphics, user experience, and sound effects.

Comparing mobile devices with other platforms, such as portable consoles, the mobile devices were not designed for playing games. This brings a set of challenges that mobile game developers need to consider while designing and developing mobile games. Korhonen and Koivisto (2006) identified the characteristics that distinct mobile games from other gaming platforms which are: small screen size, insufficient audio capabilities, limited processing power, and battery limitations. These characteristics need to be translated into requirements for developing mobile games (Korhonen & Koivisto, 2006).

Although mobile games have brought challenges to the game development community, it has also created new possibilities of making new types of games by using the sensors that mobile devices have, such as location-based games, augmented reality games, among others (Korhonen, 2016).

2.2.1 Games Components

According to Clanton (1998, as cited in Soomro, Wan Ahmad, & Sulaiman, 2013), the human-computer interaction of a game can be categorized into three groups: Game Interface, Game Mechanics, and Gameplay.

2.2.2 Game Interface

One of the elements of a video game is the interface. Korhonen (2016) mentions that the game interface is the layer that comprises the visible and audible elements of the game. It mediates the communication between the internal mechanics of the game and the player, which in the end creates an engaging player experience through its gameplay and storytelling elements (Korhonen, 2016).

The interface fulfils two aspects of the game: interaction model and perspective. The interaction model defines the way the player interacts and behaves in the digital world. The game interface shows to the players the results of using the commands of the game, which are converted into actions that the digital character performs in the digital world. On the other hand, the perspective defines how the player sees the digital world and the way the camera of the game is located. (Adams & Rollings, 2007, as cited in Korhonen, 2016)

The game presents the results of performing an action inside the game through visible and audible ways (Adams & Rollings, 2007, cited in Korhonen, 2016). The challenge of designing a game interface is to select what kind of data is presented in the interface when the player performs certain actions and what kind of data is shown to the user when the game state changes (Schell,2008, as cited in Korhonen, 2016).

2.2.3 Gameplay

Gameplay is a word that is present in many different areas of the game industry, from the game designers to the game users. The word gameplay (also written as game play) is used to describe the set of rules, challenges, and plot that the user needs to follow to complete the game (Matshafeni, 2017). The gameplay is an aspect that describes the things that you can do in the game, also known as the "what you can do" factor (Fabricatore, 2007). According to research from Fabricatore Carlo (2007) called what is gameplay according to players, one of the most important elements that define the quality of a video game is its gameplay. Gameplay is an important factor to be taken into account to provide a good experience to the player and is necessary for a game to have good playability (Riwinoto & Pertiwi, 2019). That is the reason why, in order to produce a good game, the game designers and game developers need to pay attention to the gameplay.

According to Guardiola (2019), gameplay consists of a group of actions that are performed by the player to complete the objectives of the games. The actions are all the intentional activities, including cognitive, that the player intentionally performs in the

game (Guardiola, 2019). For example, in a platform game, the user needs to pass through different obstacles to reach their goal. Among the actions that are part of the gameplay are jumping and crouching.

Also, another part of the gameplay consists of the interaction between the player and the game components. The game components are the elements that form a game, such as the game rules, the game world, and the game objects. By analysing how the player interacts with the game components, it helps to evaluate the gameplay. (Guardiola 2019)

During the design of a game, the gameplay is defined by the actions the player needs to perform during the game. This can be visualized using chart flows that represent the player's actions. The actions that are presented in the game need to provide a certain amount of challenge to deliver good gameplay. (Guardiola, 2019)

One of the aspects of gameplay that is harder to document is the emotions of the players. However, the emotions of the player are considered by the developers while creating the gameplay. To evaluate the impact of the game on the player's emotion, can be organized playtests and focus groups where they ask questions about the game experience. (Guardiola 2019)

2.2.4 Game Mechanics

To understand the gameplay of a game, it is required to know the game mechanics (Sicart, 2008). The game mechanics is defined as the system of rules that defines how the player is going to interact with the game (Sicart, 2008). The game mechanics comprises the rules that describe how the player will interact in the digital world (Adams, 2003, cited in Korhonen & Koivisto,2006).

Fabricatore (2007) describes the game mechanics as "black boxes" that receive inputs and produce outputs. This changes the inner state of the black box, which at the same time impacts other black boxes as a chain reaction (Fabricatore, 2007). The game mechanics consists of the individual rules that regulate how the player is going to interact with the virtual game. Each black box is a rule that reacts to the player's interaction and changes its state (Fabricatore, 2007). The game mechanics are one of the main components of a game design document since it states how the player is going to interact with the virtual world and how that will have an effect on the gamer experience (Sicart, 2008).

One element that determines the mechanics of a game is the goals that the player needs to achieve to complete the game. The goals of the game define the actions that the player needs to perform. However, some games allow the player to explore the virtual world without any specific goal to achieve. For example, the game Crackdown is an open-world game where the player can freely explore the virtual world without any specific goal). In this example, the game mechanics of the game are determined by the interactions between the player and the virtual environment. Although the goals of the game help to determine the mechanics of the game, they are not mandatory to be present. (Sicart, 2008)

The game mechanics of a game can be analysed using the MDA framework. This framework is used to analyse the relationship between the algorithmic elements of a game and the way the players can interact and manipulate them. This framework defines mechanics as the group of actions, behaviours, and control methods that can be performed, and experience inside a game. (Sicart, 2008)

Another approach that can be used for analysing and defining the mechanics of a game is through the object-oriented framework. The object-oriented framework is used for computer programmers to design software using objects to represent the elements of the software. Although this framework is commonly used by developers for designing and creating software, it can also be used by game developers to design the mechanics of the game. Through this approach, the game mechanics can be considered as methods that are invoked by the game elements to alter the state of the game. By using this approach, it is possible to analyse and design game mechanics more analytically and formally, for example by using Unified Modelling Language (UML) diagrams. (Sicart, 2008)

By using an object-oriented perspective, the game mechanics can be analysed as methods that describe the actions that the gamer will do, which are affected by the rules of the game (Sicart, 2008). For example, in the game Battlefield 5(a first-person shooting game), the game mechanics that can be identified are: shoot, reload, aim, change of weapon, change of position, communicate with players, sprint and crouch. These actions are limited by the rules of the game (Sicart, 2008). For example, the player can communicate with only other players of the same team.

Within the game industry, the mechanics of a game can be categorized depending on its importance inside the game. They use terms such as primary and secondary mechanics. (Sicart, 2008)

The primary game mechanics, also known as core mechanics, is defined as the actions that the players do on repetitive occasions throughout the game in the play context (Sicart, 2008). There are actions that the player performs repeatedly during the game, such as starting the game. However, these actions are not part of the primary game mechanics because they are not involved in the game context (Sicart 2008). For example, the primary mechanics of the game battlefield 5 are: shooting, sprinting, reloading, change of weapon, move the character, aim, and crouch. These are the mechanics that are the most relevant for the player to achieve the objectives of the game. Sicart (2008) mentions that the primary mechanics are usually mapped to the controller input system so that players can start interacting with the virtual world.

While the primary mechanics are required most of the time by the player to complete the objectives of the game, the secondary mechanics are used from time to time (Sicart, 2008). Sicart (2008) states that the purpose of the secondary mechanics is to support the players to play the game. For example, selecting and customizing the character of a game. Although customizing the character of the game is part of the game experience, the customization of the character can be omitted, and the player would still be able to play the game.

The differentiation between primary mechanics and secondary mechanics helps to understand the different mechanics of the game and the impact that it has on the game experience. This allows for a better understanding and design of the way the game is going to work.

2.3 Usability

Usability is defined by Cambridge University as "the fact of something being easy to use, or the degree to which it is easy to use" (Cambridge Dictionary n.d.). This is applied to any product that is used by a human. Usability is also defined by Goodwin (1987) as "the extent to which a system and its users communicate clearly without misleading or

misunderstanding" (as cited in Abdulhak et al., 2011, p. 7). Nielsen (as cited in Abdulhak et al, 2011) describes usability as the quality of the experience of the user when using any software through an interface.

There are diverse methodologies that are used to evaluate the usability of a product. One of the usability methodologies that is used the most by the community is heuristic evaluation. This evaluation consists of a group of specialists that evaluate the usability of a certain product by comparing the product with a list of good usability principles. The result of a heuristic evaluation is a list of usability problems that are caused due to violations of usability principles. (Mylly et al. 2019)

2.4 Game Usability

The importance of usability over the years has increased. It is even considered in game development good usability is not just an advantage, but it is necessary to compete with other games (Mylly et al., 2019). Furthermore, the usability of a game and the quality of its interface are aspects that players care about while choosing which game to play (Rajanen & Marghescu, 2006).

Game usability refers to the use of usability methods and principles in the gaming context (Korhonen, 2016). Pinelle et al. (2008) defined game usability as "the degree to which a player is able to learn, control, and understand a game". Game usability also analyses the interaction between the players and the digital world (Soomro et al., 2013). According to Rajanen and Nissinen (2015), the term of game usability is viewed from two different perspectives: the first one evaluates game usability based on its interface, controls, and other aspects. Contrary, the other group studies game usability based on its gameplay, game mechanics, and game interface (Rajanen & Nissinen, 2015). Also, it includes other more abstract elements such as the degree of enjoyment and fun (Rajanen & Nissinen, 2015).

According to Alhaidary & Altammami (2017), three aspects are taken into account while evaluating the usability of a game: game interface, gameplay, and game mechanics. Alhaidary & Altammami also states that the game usability is evaluated using a wide range of methodologies, which includes interviews, questionnaires, or observations. The most common methodology for evaluating the usability of a game is heuristic evaluation (Alhaidary & Altammami, 2017).

For a game to be successful it requires the gamer to be able to go through the game without any obstacle. Therefore, the game needs to provide a good game experience. To offer a satisfactory game experience, it requires the game to have a proper user interface that allows the player to be able to focus mainly on the game. (Korhonen, 2016)

Another important aspect that affects the usability of the game is the game design. If the game is too hard to play it affects the game experience of the player and it will discourage the gamer to continue playing since it becomes frustrating to play. Therefore, at some point, the gamer will stop playing. On the other hand, designing a game that is too easy to play creates a negative effect on the user experience. Without challenging the gamer to develop their skills due to its low entrance level, it will not appeal to the gamer to keep playing since it is boring. Achieving the balance of challenge and ease to use is one of the aspects that need to be considered while applying usability in the design of the gameplay of a game. (Korhonen, 2016)

The usability of games is important for the success of a game, especially nowadays where the amount of competition and new games are increasing. Among the reasons the game usability is important are: first, it allows to find small glitches in the interface of the game, the presence of glitches causes a bad and annoying experience for the game players. Another reason for adopting game usability practices is that it differentiates the game compared with other games where game usability has not been considered. With time games are becoming more complex, which increases the possibility of encountering errors related to the game interface or game mechanics. (Daud et al., 2016)

For a deeper analysis of the interaction between the player and the game and all the different attributes that are involved, the concept of playability was created. In the next section more information about playability is presented.

2.5 Playability

At the beginning of gaming history, it increased the interest in using usability practices to improve the quality of the game. However, it is not possible to apply classical usability testing to evaluate game usability because of the differences between a game with other kinds of products (Sánchez et al., 2012). For example, one of the attributes of a good user experience is to make the tasks of the users easy and quick to complete, while one of the main objectives of games is to make the user experience challenging so that it retains the interest of the player (Soomro et al., 2013). Due to the different purposes a game has compared to other kinds of software, it is not possible to use only usability principles in the video game context (Soomro et al., 2013). Table 1 presents the differences between usability goals and playability.

Table 1. Difference between usability and playability. Reproduced from Sánchez et al. (2009), p. 66 cited in Novick (2014)

UX Usability Goals: Productivity	PX Playability Goals: Entertainment
Task completion	Entertainment
Eliminate errors	Fun to beat obstacles
External reward	Intrinsic reward
Outcome-based rewards	Process is its own reward
Intuitive	New things to learn
Reduce workload	Increase workload
Assumes technology need to be humanized	Assumes humans need to be challenged

The word playability is used to describe the properties of a videogame, such as its game mechanics, its rules, and its design (Sánchez et al., 2012). Sánchez's (2009) definition of playability is "playability represents the degree to which specified users can achieve specified goals with effectiveness, efficiency and especially satisfaction and fun in a playable context of use" (p. 1034). Playability is a framework that is used to evaluate

games; therefore, the player experience of the game can be measured and improved (Sánchez et al., 2012). The concept of playability is related to the term's fun, fulfilment, flow, satisfaction, engagement, player experience, and pleasure (Wong, Chu, Khong, Lim, 2010). According to a study related to playability on video games, playability is defined as a list of properties that describes the player experience while using a game system whose main purpose is to entertain the user while playing alone or in a group for the purpose of evaluating the player experience (Sánchez et al., 2009, as cited in Wong et al.,2010). In both definitions of playability, the evaluation of the experience of the player while gaming and the degree of fun a player experiences are important elements of playability.

One of the most complete definitions of playability can be found in the Usability Glossary (2002, as cited in Korhonen, 2016) which describes playability the following way:

> The degree to which a game is fun to play and usable, with an emphasis on the interaction style and plot-quality of the game; the quality of gameplay. Playability is affected by the quality of the storyline, responsiveness, pace, usability, customizability control, intensity of interaction, intricacy, and strategy, as well as the degree of realism and the quality of graphics and sound (p. 22 – 23).

As it can be observed from the previous definition of playability, there are many elements that are involved while evaluating this phenomenon. This definition identifies the main elements that conforms playability.

Playability is used to assess the interaction between the players and the game (Korhonen, 2016). By doing so it helps the game designers to have a better understanding of the perception of the gamer related to the game (Ponnada & Kannan, 2012). This allows the game designers to find an area of opportunity that can improve the game experience of the player (Ponnada & Kannan, 2012). Playability is used to measure the positive or negative impact a game has produced on the player experience (Sánchez et al., 2012). With the help of playability, it helps to quantify and detect the usability attributes that a game has (Sánchez et al., 2012). The concept of playability is used as a guideline on how to create a game that has good gameplay and provides social entertainment (Korhonen, 2016).

Playability problems arise usually when the elements of the game are confusing, or it is not clear the way the player needs to interact with the game. Therefore, it is hard for the player to follow the game. Hence, all the aspects of the games need to be evaluated in a playability evaluation. (Korhonen, 2016)

Another aspect of playability is that it helps to identify the elements of a video game related to user experience and identify the impact it has on different players (González Sánchez et al., 2012).

Playability is a set of factors that influences how fun or entertaining a game can be (Somasundaram, 2012). Some of the factors that affect the playability of a game are quality of the storyline, responsiveness, pace, usability, possibility to customize, control, intricacy, the intensity of interaction, strategy, degree of realism, quality of graphics, and quality of sound (Sánchez et al., 2012).

According to Korhonen (2016), his definition of good playability is:

A game has good playability when the user interface is intuitive and the gaming platform is unobtrusive, so that the player can concentrate on playing the game. Fun and challenge are created through gameplay when it is understandable, suitably difficult and engaging (p. 35).

As it can be observed in the definition of Korhonen about playability, fun and challenging games that are easy to understand are the goals that games need to have in order to achieve good playability. Another way to measure the playability of a game is by measuring the degree of immersion players have while playing a game. This can be observed when the user is in a state of flow during the gaming session (Sánchez et al., 2012).

2.5.1 Playability Attributes

Although playability is based on the fundamentals of usability, there are other characteristics and elements that are specific to the video game context, such as the degree of entertainment that brings to the user (Sánchez, et al., 2012).

In a research written by Sánchez et al. (2012) about playability, it was proposed seven characteristics that can be used to measure playability, which are: satisfaction, learnability, effectiveness, motivation, emotion, socialization, and immersion. These properties are used to assess the user experience of the game and the interaction of the user with the game.

Satisfaction

The purpose of measuring the satisfaction of a game is to understand how fun is to play a game. This property is highly subjective since each person has different notions of what is a satisfying experience. But in the game context, the degree of satisfaction can be measured by evaluating how entertaining a game is. (Sánchez et al., 2012)

To measure the satisfaction of a game, the following properties can be used to detect it. One of them is fun. By observing if a game is fun to play, it produces a satisfactory experience for the gamer (Sánchez et al., 2012).

Another aspect that can be used to detect the satisfaction of a game is through the player's disappointment. When a player is disappointed in the experience of the game, it indicates that the user is not satisfied with the game. This is a property that is important to consider for producing games that have good playability. (Sánchez et al., 2012)

The last property that is used to measure the satisfaction of a player over a game is attractiveness. If a game is appealing to the user, it makes the user keep playing (Sánchez et al., 2012).

Learnability

Learnability is the player's ability to understand and learn the objectives and rules of the game. Although in other types of software applications the aim is to reduce the learning effort to use the program. In the gaming context, it is possible to change the learnability according to the gameplay and game design. For example, the learning curve of the game can be high at the beginning to make sure that the players have all the required knowledge to play the game. Another approach can be that the player gets only the required

knowledge to complete the level, but new skills are achieved throughout the game. (Sánchez et al., 2012)

Effectiveness

In the context of video games effectiveness is measured by the amount of time and resources it requires to create an entertaining experience while the user goes through the objectives of the game. A video game that can be considered effective is able to attract a player's attention from the beginning till the end of the game. Among the elements that are contained in the effectiveness of a game are completion and structuring. (Sánchez et al., 2012)

Completion is used to measure the effectiveness of a game. The reason for this is because it allows the players to understand the mechanics of the game and the goals that need to be achieved. If the players can complete the game, then the game is effective. One aspect that might influence the completion of the game is the player profile. For example, casual gamers might complete the primary challenges of the game, but they might omit the secondary challenges since they are not interested, and they do not play that much. On the other hand, hardcore gamers get more involved in the game and they might complete all the challenges of the game. (Sánchez et al., 2012)

The other element that describes the effectiveness of a game is its structure. The structuring of the game refers to how the elements and objectives are presented during the game. A properly structured game presents the challenges and objectives in a balanced way that keep the players engaged and keeps them interested to complete the game. (Sánchez et al., 2012)

Immersion

This attribute is used to detect the degree of involvement that a user has while playing the game. When a game can be considered to achieve a good level of immersion occurs when the player can interact properly with the game and can achieve the game objectives (Sánchez et al., 2012). In the research paper written by Sánchez et al. (2012) they identify 7 properties that characterize the immersion of a game: conscious awareness, absorption, realism, dexterity, interactive dexterity, virtual dexterity, and social-cultural proximity.

The property conscious awareness refers to how the player is aware of the effects of his actions in the virtual world. By recognizing the consequences of his actions in the game allows the player to understand the gameplay and be able to complete the game. (Sánchez et al., 2012)

To detect and measure the amount of attention a player has while playing, the absorption attribute is used. It is possible to identify when a player has a high degree of absorption since the player is focusing all the attention on the game. (Sánchez et al., 2012)

Another aspect that affects the immersion of the game is its realism. With realism, it does not mean that the graphics and sounds do not have to be close to real life, but they have to be good enough that they are believable for the player. Games with a high level of realism tend to be more attractive for the gamers and therefore achieve a high level of immersion. Thus, the player can focus on the rules and the objectives of the game. (Sánchez et al., 2012)

One More attribute that is important to consider while measuring the immersion of a game is dexterity. This attribute is used to understand how the player can control the game using different kinds of movements. If the game is not responsive when the player is moving the game controls, it will have a negative impact on the game immersion. (Sánchez et al., 2012)

Additionally, to achieve immersion needs to be taken into account the social-cultural proximity. When the elements that are contained in a game are close to the player's social-cultural background, it will make the player feel more identified with the game and it will appeal to the gamer to continue playing the game. For example, the language that is used for games that are targeted at an adult population is different from the language that is used for games for children. (Sánchez et al., 2012)

Emotion

While a person is playing a game, there should be different reactions and feelings that need to be involved, depending on the genre and type of game. If the game can trigger different kinds of emotions, it means it provides a good player experience. With emotion, refers to all the involuntary stimulus that the video game produces. The following properties are related to the emotions while playing a game: reactions, conduct, and sensory appeal. (Sánchez et al., 2012)

During a game session, different stimuli are presented to the user, this makes the user react. There are four types of reactions: internal, behavioral, cognitive, and social. The internal reaction is caused automatically by the human senses. The behavioral reaction is presented when the gamer interacts with the virtual world and produces certain reactions, such as happiness or surprise. The cognitive reaction occurs by the thoughts and memories the gamer has while playing the game. Social reactions arise when the player is interacting with other players in the game. Playing with somebody triggers different reactions than playing alone. (Sánchez et al., 2012)

Motivation

For the purpose of making sure that a player is able to complete the game, there should be a motivation to do it. Motivation is the characteristic that is present in the game that makes the players keep playing the game and complete all the objectives. For achieving a high degree of motivation inside a game it should be encouraged to improve the skills of the player and there should be positive feedback when the player is able to reach the game's goals. The next attributes are observed in the motivation of the game: encouragement, curiosity, self-improvement, and diversity. (Sánchez et al., 2012)

For the players to feel motivated to keep playing the game they need to be encouraged. For example, if a player is facing a challenging objective, a high reward needs to be given to the player to motivate the player to continue with the next level. If the amount of reward is low in comparison with the amount of time and effort the player needed to put in while completing the level, it will cause frustration over the player and discourage the player to complete the next game objective. (Sánchez et al., 2012)

Another property that is involved in the motivation of the player is curiosity. To make the gamer keep playing, it is necessary to provide new elements during the game that will make the player explore and discover new things. (Sánchez et al., 2012)

One More property that affects the motivation of the player is the self-improvement of the player. This is an important characteristic since the players are looking for improving their skills to dominate the game better. Without challenging the gamer and making the player sharpen their skills, they will make the players bored and even stop playing. (Sánchez et al., 2012)

Diversity is an important attribute that affects the motivation of the player. There is a relationship between the motivation of the player and the diversity of elements in the game. Having different elements inside the game makes the game more interesting for the player. For example, if the game has levels that look very similar and there are not that many new elements that differentiate from the previous level, it will make the gamer feel bored and therefore quit the game. (Sánchez et al., 2012)

Socialization

An additional aspect that is involved in the playability of a game is the social factor. This factor allows the researchers to analyze the interaction of players inside the game. The experience the gamers have when they played by themselves is different from the experience they have when they plate with other gamers. Another aspect of socialization is involved in the way the player personalizes his character in the game. (Sánchez et al., 2012)

2.5.2 Playability Evaluation Methodologies

Heuristic Evaluation System

For an organization to stay competitive in the market, it is mandatory to have tools that allows to evaluate the playability and user experience of video games (Sánchez et al., 2012). One of the methodologies that are used to evaluate how the effectiveness and efficacy of video games is heuristic evaluation. This methodology was first used by Nielsen and Molich (Soomro et al., 2013). The motivation of this methodology is to inspect and test the user interface of software against a list of rules that the software should comply with to pass the evaluation (Soomro et al., 2013). During the initial stages of heuristic evaluation, it was used only to inspect the user interface of a software product (Soomro et al., 2013). But later it was used as guidelines during the development and design of the software (Soomro et al., 2013). Thus, at the moment of testing the usability of the product, it will pass the heuristic test (Soomro et al., 2013).

During the 1990s this methodology was used only on evaluating the design of the interface of a software. However lately this methodology has been used by other industries to evaluate the usability of a product, for example, the gaming industry. (Soomro et al., 2013)

According to Nielsen (1994, as cited in Alhaidary & Altammami, 2017), heuristic evaluation consists of a small group of evaluators that inspects the interface and evaluates the degree of compliance of the interface based on a set of usability principles known as "heuristics". The first heuristics were developed by Nielsen and Molich and the purpose of the heuristic was to evaluate the interface of software (Soomro et al., 2013). Due to the difference between video games and other kinds of software programs, a new set of heuristics were introduced by Clanton in his study about how computer game can be designed to engage users (Soomro et al., 2013). In 2009, a different set of heuristics for

evaluating video games called PLAY was introduced by Desurvire (Soomro et al., 2013). This heuristic comprises three categories: the first category evaluates the gameplay of the game, such as the goals, the challenges, and the controls (Aker, Rizvanoglu, Inal, & Yilmaz, 2016). The second category assesses the degree of entertainment and the emotions that are produced while playing the game (Aker et al., 2016). The last category evaluates the usability and game mechanics of the game, for example, the screen layout, the navigation, the game story immersion, among others (Aker et al., 2016). Then, in 2006, Hannu Korhonen and Elina Koivisto presented heuristics which were designed for evaluating mainly mobile games, check Appendix B (Aker et al., 2016).

Table 2. Adaptation of diagram found in Aker et al., (2016) about the elements that are evaluated using the proposed heuristic evaluation based on the author

Sánchez et al.	Federoff	Desurvire et al.	Korhonen & Koivisto
• Artistic Playability • Intrinsic Playability • Mechanical Playability • Interactive Playability • Interpersonal Playability • Intrapersonal Playability	• Game Play • Mechanics • Game Interface	• Game Story • Game Play • Mechanics	• Game Play • Game Usability • Mobility

Due to this difference between traditional software products and games, heuristic evaluations needed to be adapted to fit the context of games. According to Soomro et al. (2013), one of these factors that differentiate a game from other software is the enjoyability factor. He states that games need to be fun and entertaining to achieve their goal for the user. Therefore, an adapted version of heuristic evaluation, known as playability heuristics, was created. In 1998 Clanton published a list of game principles that were used to inspect formally and informally the usability of a game (Soomro et al., 2013). The purpose of this list is to understand how to design a computer game that can be engaging to the user (Soomro et al., 2013). After that, in 2004, it was created a new proposal for inspecting games called "The Heuristics for Evaluating Playability" (Soomro et al., 2013). Later, in 2006 a new evaluation for inspecting games was created by Hannu Korhonen and Elina Koivisto (Soomro et al., 2013). Table 2 presents the different elements that are taken into account by each author for evaluating the playability of mobile games.

Playtesting Game

Another methodology that is used to evaluate the playability of a game is through a playtesting session. A playtesting session consists of a player who is invited to a laboratory where the test is conducted. During this session, the participant is asked to play a game and perform certain actions inside the game. At the end of the session, questions are asked related to the experience of the player, graphics, story, sound, controls, among

others. The player is observed during the session to analyze their reactions towards the game. The playtesting sessions are performed systematically and used scientific methods to observe the session. (Davis, Steury, & Pagulayan, 2005)

The goal of the playtest is to evaluate the gameplay and the player experience of the analyzed game. Also, they are used to detect bugs and gaming glitches that can occur during the game. Glitches are software errors that cause unexpected behavior in the game. (testbytes, 2019)

The playtests are performed through the development of the games, from the game design to the testing of the game before releasing the game. The first playtest known as gross playtesting is conducted during the design phase of the game. The focus of this stage is to find faults in the gameplay, usually, this test is conducted by game designers. (testbytes, 2019)

When the game is still in development, it is carried out as an in-house playtesting. This session aims to find bugs and glitches and keep testing the gameplay and game mechanics. This test is conducted by contracted play testers. The blind test is conducted by testers who do not have any experience with the game before the test. These are used to have new insights that were not detected in the previous tests. (testbytes, 2019)

The final testing occurs in the last phase of the game development, where the final version of the game is tested. Few improvements are made before publishing the game. (testbytes, 2019)

Think-aloud test

Think aloud test is a commonly used methodology for evaluating the usability of a product (Nørgaard & Hornbæk, 2006). Nielsen (2012) defines think aloud test "in a thinking aloud test, you ask test participants to use the system while continuously thinking out loud, that is, simply verbalizing their thoughts as they move through the user interface". This allows the evaluator to understand what is going through the user's mind while using a product.

One of the benefits of using this methodology is its simplicity of implementation. It is possible to conduct a think-aloud test without the need of buying equipment or the need for special facilities to conduct the test. Another advantage of using this evaluation method is its flexibility to use. It is possible to apply this methodology at any phase of the development life cycle, including the early stages where a paper prototype can be tested. This methodology is suitable for projects that use agile practices. (Nielsen, 2012)

On the other hand, there are certain threads of using the think-aloud test. One of them is that it can introduce biased in the behavior of the user. During the test, the evaluator might introduce thoughts that do not belong to the user. Another problem with this methodology is that it is hard for the user to talk and use a product at the same time. This makes it harder for the user to give valuable feedback on the product that is being tested. (Nielsen, 2012)

2.6 User experience (UX) and Player experience (PX)

2.6.1 User experience (UX)

As it was mentioned in the previous chapter related to usability, the purpose of usability is to study the interaction between the user with the product. Organizations need to develop products that provide a nice experience while using them. To study and measure the experience of a subject while using a product or service, the term user experience was used (Sánchez et al., 2012).

User experience goes beyond usability (Sánchez et al., 2012). While usability focuses on the effective and efficient interaction of the user with the object, user experience focuses more on evaluating how the users feel while interacting with the object (Sánchez et al., 2012). Bevan (as cited in Rajanen & Rajanen, 2017) mentions that the main objective of user experience is to analyze user preferences, perceptions, emotions, and physical and physiological responses of users towards a product.

2.6.2 Player experience (PX)

Due to the different attributes and objectives that a game has compared to other kinds of products, a more precise term was needed to describe the experience of the gamer. Instead of using the user experience in the gaming context, the player experience was conceived. Although the term game experience is also used to describe the experience of the player, the term player experience is a more adequate and precise term since it is used to describe the experience of the player. (Olsen, Procci, & Bowers, 2011)

The difference between playability and player experience is the subject that is under evaluation. While in playability the focus of the study is the game, the focus of study of player experience is the player, see Figure 1. (Nacke, 2009)
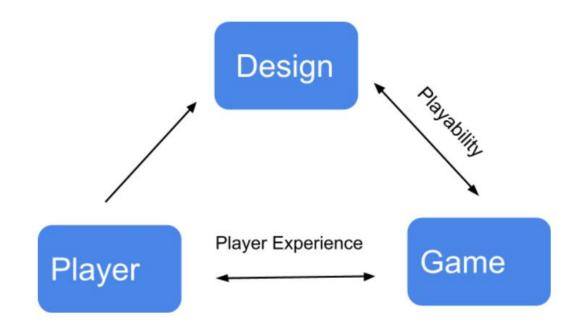

Figure 1. Adaptation of diagram found in Nacke et al. (2009) shows the elements that are involved during the analysis of player experience vs playability

To measure and understand the player experience of an individual, there are used different methodologies to evaluate this phenomenon. The factors that are studied while analysing the player experience are gameplay performance, human emotions, and psychological characteristics (Olsen et al., 2011).

Some of the methods that are used to evaluate the player experience are psychological player testing, eye tracking, persona modelling, game metrics, and qualitative interviews. These methods can be categorized into three groups: physiological methods, psychological methods, and behavioural methods. (Olsen et al., 2011)

The physiological methods analyse how the body reacts according to the stimulus that the game produces (Nacke, 2009). To conduct these methods, sensors are placed in the individual's body and measure changes in the body. Among the tests that are used to evaluate the game experience through the reaction of the body are electroencephalography (EEG), electromyography (EMG), and electrodermal activity (EDA) (Olsen et al., 2011). By understanding the effect the game has on the player it is possible to have a better understanding of what the player is experiencing. With these measurements, it is possible to infer the emotions of the player. Among the benefits of using physiological evaluation is that it is feasible to get data without the need of interfering with the player (Nacke, 2009).

The psychological methods are used to assess the feelings and emotions of the individual after playing a game. One of the most common methods that are used is through surveys. For example, the Game Experience Questionnaire (GEQ). is an evaluation that consists of 36 different questions that are used to evaluate the competence, immersion, flow, tension, challenge, engagement, positive, and negative effect of a game.. (Olsen et al., 2011)

According to Korhonen (2016), a game can be considered to have a good player experience when the gamer is able to interact with the game world and, if applied, with other players in an effortless and unobtrusive way. In other words, the player can play the game without any issue. A game with good player experience is considered to be balanced and playable (Korhonen, 2016).

3.   Systematic Literature Review

The purpose of this chapter is to introduce the concept of SLR and to explain how SLR was performed in this thesis. First, section 3.1 provides background information about the fundamentals of SLR, including the steps that need to follow to carry an SLR. Then section 3.2 describes the reason why an SLR was conducted to research this topic. Next, section 3.4 depicts the planning phase of the SLR. Finally, section 3.4, presents how the actual SLR was performed and reported.

3.1 Fundamentals of Systematic Literature Review

A systematic Literature Review consists of three main phases: Planning the review, conducting the review, and reporting a review.

In the first phase, planning the review, it is identified the need for conducting an SLR and it is developed a protocol that will be used to identify the studies that will be utilized in the research (Kitchenham, 2004).

The second phase of an SLR is the stage where the actual SLR is conducted. During this phase, the primary studies are identified and selected. After that, the quality of the selected studies is evaluated. Then the data extraction is performed. (Kitchenham, 2004)

The final stage of an SLR consists of reporting the findings of the research. Although the phases seem to be performed in sequential order, some phases are required to be iterated (Kitchenham, 2004). For example, the development of the research protocol is refined through iterations (Kitchenham, 2004). The steps for performing an SLR are summarized in Figure 2.
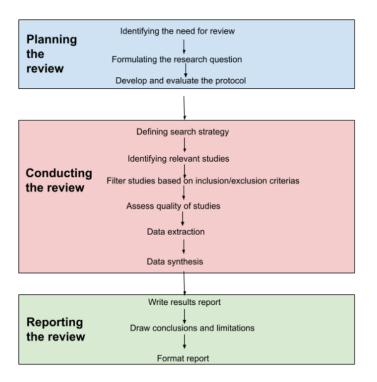

Figure 2. Diagram that shows the phases for conducting a SLR

3.2 Reasons for Performing a SLR

As it is mentioned by Kitchenham (2004) among the most common reasons for conducting an SLR are:

- To summarize empirical studies of a specific topic
- To identify any gaps in the current body of knowledge that requires further research
- To provide a background that will promote further research activities

The purpose of this study is to identify what are the methodologies that are used to evaluate the playability of video games on mobile devices and if there is a recommended methodology for evaluation mobile games. Due to these reasons, the most appropriate research method that will allow finding the answers to this investigation is by conducting a Systematic Literature Review.

3.3 Planning the review

As it was mentioned in the previous section, the first step that is performed during an SLR is planning the review. According to the guideline for conducting an SLR proposed by Kitchenham (2004), the steps that comprise this phase are:

- Identification of the need for a Systematic Literature Review
- Defining the research questions
- Defining the review protocol

These steps were used in the planification of this SLR. The following section contains information about how this SLR was planned.

3.3.1 Identification of the need for a Systematic Review

According to Kitchenham (2004), the need for a systematic literature review occurs when it is required to summarize all the available information about a specific topic in an unbiased way. The purpose of this is to draw general conclusions about the studied subject. Furthermore, the SLR can also be used as a foundation for further research (Kitchenham & Charters, 2007).

In this research, the need for performing an SLR is to summarize the existing information related to playability and to identify the methodologies that are used to assess the playability of mobile games. To achieve this, it is required to detect and extract all the information related to playability in the mobile game context.

3.3.2 Research Question

The research questions that are intended to be answered through SLR are the following:

- RQ1: What are the methodologies that are used to measure the playability of mobile games?
- RQ2: Is there a recommended methodology for evaluating the playability of mobile games?

The purpose of the first question (RQ1) is to identify the methodologies that are used to assess the playability of mobile games. The objective of the second question (RQ2) is to identify the methodologies that are most used to evaluate the playability of mobile games.

3.3.3 Data sources

To get the data that will be used to find answers for the research questions, it will be used databases that contain studies that are relevant to the current study.

The databases that will be used for this research are:

- IEEE Xplore
- Scopus
- Web of Science

The reason for selecting the following databases was because they are scientific databases that contain research related to the field of software development and human-computer interaction.

3.3.4 Pilot search

According to Kitchenham (2004), it is recommended to perform pilot research before conducting the actual SLR. The purpose of the pilot search is to identify the relevance of the topic and to get a general idea of the number of available studies (Kitchenham, 2004). Therefore, a pilot search was performed using Google Scholar on the 15 of January 2021. By conducting a query in Google Scholar using the keyword playability only without quotes showed 20,500 results. Since the focus of this study is related to understanding playability in the video game context, the search string was changed to "playability in video games" without quotation marks, which showed 10,400 results. By adding quotation marks in the keyword playability, it threw 18,700 results. Since the thesis aims to understand the playability of mobile games, was performed a search using the string without quotation marks "playability in mobile games", which showed 6,010 results. By searching with quotation marks using the keyword "playability in video games" resulted in 35 results. Further search using a combination of two keywords using quotation marks "playability in video games" and "usability in video games" with the logical operation OR was done, which resulted in 61 results.

3.3.5 Developing Review Protocol

One of the main elements that need to be defined before performing an SLR is to specify and review the protocol. This protocol mentions the methods that will be used to perform an SLR (Kitchenham, 2004). The purpose of the protocols is to reduce the possibility of researcher bias (Kitchenham, 2004). According to the guideline for conducting an SLR proposed by Kitchenham (2004) the elements that are included in the review protocol are:

- Background of the study
- Research questions that are answered through the review
- The search strategy that will be used for identifying the primary studies, such as search terms and the resources that will be utilized (databases, journals)

- The criteria that will be used for selecting the primary studies, also known as inclusion/exclusion criteria
- A checklist that is used for assessing the quality of the studies
- The data extracting strategy that determines how the information from the primary studies is found
- The synthesis strategy that defines how the results are summarized and presented

The protocol was submitted for review and approval to the supervisor before conducting the SLR. The protocol that was used for this research is described in the following sections.

3.3.6 Search Strategy

The main objective of an SLR is to find as many studies as possible related to the topic that is being studied. For this purpose, it is required to develop a search strategy that will help to achieve this goal. This process is iterative where different search strings are applied to databases with the help of Boolean operators AND's and OR's. (Kitchenham, 2004)

As part of developing the search strategy, it is recommended by Kitchenham (2004) to break down the questions into individual elements and create a list of possible synonyms and abbreviations. Table 3 shows the elements and possible synonyms that are used for constructing the search string.

Table 3. Shows the elements that form the search string

Concept 1	Concept 2	Concept 3
• Playability • Usability	• Video game • Video-game • Videogame • Digital game • Electronic game	• Mobile • Smartphone

After that, it is feasible to construct a sophisticated search string to find as many materials as possible that will allow us to answer the research questions. In this SLR, with the help of the pilot search, it was possible to identify the elements that will be part of the final search string.

The final search string developed after an iterative process was:

(("playability" OR "usability" AND NOT "accessibility") AND ("video games" OR "video-games" OR "video game" OR "video-game" OR "gam*" OR "digital game" OR "digital game*" OR "electronic game*" AND NOT "gamification") AND ("mobile" OR "phone" OR "smartphone"))

3.3.7 Selection Criteria

To find the relevant primary studies that will be used to answer the research questions, it is required to apply certain criteria. These criteria will guide the researcher to conduct an

SLR with as few biases as possible. The selection criteria are usually defined during the definition of the protocol. These criteria should be established based on the research questions. (Kitchenham, 2004)

The inclusion criteria that were used for selecting the primary studies are the following:

- Include if Paper contains topics related to playability or usability of mobile games
- Include if Paper contains information related to playability (usability in gaming context) methodologies in the context of mobile games
- Include if Paper focus is to study the playability of native mobile games
- Include if Written in English
- Include if Paper is peer-reviewed
- Include if Full paper is available

The exclusion criteria's that were used for selecting the primary studies are the following:

- Exclude if Paper is not related to playability of mobile games
- Exclude if Paper only mentions playability of mobile games but is not the main topic
- Exclude if Paper mentions non-digital games
- Exclude if Paper is not about playability or usability in mobile games
- Exclude if Paper focus of study is using the mobile device as an input controller
- Exclude if Paper focus of study is a VR game or a game that uses wearable devices
- Exclude if Paper focus of study is a web base mobile game
- Exclude if Paper uses wearable devices to play the game
- Exclude if App is not a video game
- Exclude if Paper studies the playability of a not functional game prototype
- Exclude if Paper is duplicated
- Exclude if Paper is not written in English
- Exclude if Paper is not peer-reviewed
- Exclude if Paper full content is not available

According to studies (Kitchenham, 2004), the inclusion and exclusion criteria are applied based on the title and abstract of the study. Kitchenham also mentions that in case it is not possible to decide if the paper is included/excluded based on its title and abstract, the full-text version of the paper needs to be retrieved and read.

3.3.8 Quality Assessment

To reducing the bias of the study, is required in an SLR to assess the quality of the primary studies (Kitchenham, 2004). According to Kitchenham (2004), among the reasons it is required to assess the quality of the studies are:

- To offer more detailed inclusion/exclusion criteria
- To assist in understanding if the difference of the results is related to the quality of the study
- To help in interpreting the results of the research and determine its validity
- To provide recommendations for additional research
- As a way of evaluating the importance of each study when the results are produced

The quality of the studies that will be selected in this research will be done by answering the quality assessment questions, check Table 4.

Table 4. Quality assessment questionnaire

Quality Assessment Question	Yes/No
Is the objective of the paper explicitly mentioned?	
Does the paper explicitly mention the research methodology?	
Does the research paper contain results?	
Does results of the study contain useful information related to the research questions?	
Does the study mention the limitations and future research of the study?	

The studies that pass the quality assessment questions are selected for the final list of primary studies used in this thesis. To perform the quality assessment, the full-text version of the study is read and then it is evaluated its quality.

3.4 Conducting the review

This section of the research explains how the SLR was conducted. This phase of the SLR consist of the following steps:

- Selection Process
- Data Extraction Strategy
- Data Synthesis
- Reporting the review

After the protocol was defined and approved by the thesis supervisor, the actual SLR was conducted. The process of conducting the SLR started on the 12th of April 2021, when the search string was applied to each database to obtain the studies of the research. In the following section, it describes in more detail the selection process.

3.4.1 Selection Process

As it was established in the protocol the first step that was conducted was to identify the papers that will be used in the study. This was done after defining the search string that would be used in each database. The final search string used in the SLR was:

(("playability" OR "usability" AND NOT "accessibility") AND ("video games" OR "video-games" OR "video game" OR "video-game" OR "gam*" OR "digital game" OR "digital game*" OR "electronic game*" AND NOT "gamification") AND ("mobile" OR "phone" OR "smartphone"))

As it was mentioned in the previous section, the search of primary documents for the thesis was conducted on the 12th of April 2021, where the search string was applied to

each database. The search string needed to be modified based on the syntax of each database. After applying the search string, a total of 1,390 papers were retrieved. To identify the most relevant studies that would be used for further analysis, filters from the database were used to exclude irrelevant studies. The goal of this process was to identify papers related to the field of computer science, human-computer interaction, and user interface. Other parameters that were considered during this step were the language of the paper (English) and its availability (full-text available). During this process, 890 papers were excluded and 500 remaining. Table 5 presents an overview of the selection process.

Table 5.   Selected primary studies of this research

Database	Number of papers	Number of papers excluded at database level	Remaining number of papers	Duplicates	Total
Web of Science	458	235	223	21	479
Scopus	659	561	98		
IEEE Xplore	273	94	179		
Total	1,390	890	500		

After obtaining the most relevant primary studies for the study, they were imported to the reference management tool RefWorks. With this tool, it was possible to find the duplicate studies and come up with the final list of 479 studies that were used in the next phase.

The next step of the SLR was to apply the inclusion and exclusion criteria that were defined in the protocol. The criteria were applied by analysing the title and abstract of each paper. This process was done in 3 phases.

In the first phase, all the papers that met the inclusion criteria of papers that talk about playability/usability in mobile devices were applied. After this was done, 183 papers were identified.

In the second phase, a more exhaustive analysis of the title and abstract of the paper was performed. If the paper was not able to include or exclude based on the title and abstract, the paper was downloaded, and the full text was read. Then the papers were filtered if it meets the rest of the inclusions and exclusions criteria defined in section 3.4.7. Based on this process 124 papers were excluded, with 59 studies left.

The final phase of the selection process was to assess the quality of the studies. To do this the full-text primary study was downloaded and read. Then, the paper was analysed and evaluated based on the quality questions that were previously defined in the protocol. From this process, 11 primary studies were not able to be accessed the full text and therefore excluded from the SLR. In the end, 48 primary studies were downloaded, and the full text was read. After that, according to the quality control criteria, 21 papers were excluded, and 27 primary studies were selected for data analysis and data extraction.

3.4.2 Data Extraction Strategy

The data extraction strategy is designed to get the most relevant information that allows answering the research questions. The following data was obtained from the selected studies:

- Information related to the primary study, such as authors, title, year, research method, and research objective
- The type of mobile device that was used to test the playability of the game
- Name of the methodologies used to evaluate the playability of mobile games for answering RQ1
- The effectiveness of the mentioned methodologies for answering RQ2

The data was extracted by reading the full-text primary study, then it was recorded in an excel sheet. It is recommended by Kitchenham & Charters (2007) to use a data extraction form to collect the data that will answer to the research questions. However, due to the nature of this study, it was decided to collect the data in an excel sheet.

3.4.3 Data Synthesis

Data synthesis is the step of an SLR where the data obtained from the primary studies is analysed and summarized. There are two approaches the synthesis can be performed, which are quantitative (meta-analysis) and descriptive(non-quantitative). The data synthesis should be included in the SLR protocol. (Kitchenham & Charters, 2007)

In the data extraction phase, it was obtained information related to the methodologies that are used to evaluate the playability of mobile games. Also, it was obtained information related to the methodologies that are the most effective for evaluating the playability of mobile games.

To present the results obtained from this SLR, it was used a descriptive synthesis. The purpose of this method according to Kitchenham (2004), is to summarize the data collected from the primary studies by using tables. Kitchenham also stated that the data contained in the tables should answer the research questions. Furthermore, he mentions that the tables should be arranged in a way that shows the similarities and differences between the primary studies. Therefore, it was considered to be appropriate the use of descriptive synthesis for summarizing the results of this research.

3.4.4 Reporting the Review

The last stage of the SLR is to report the results. According to Kitchenham & Charters (2007) in this phase of an SLR, the results are written and distributed to the interested parties. For this research, the report was written based on University of Oulu Master thesis guidelines. The finalized version of this research is published in the University of Oulu Library system Laturi. The results of the SLR are presented in chapter 4.

4.   Results

This chapter contains the results that were obtained from the SLR. From the initial selection of 479 studies, 27 primary studies were selected and included in this study based on the inclusion/exclusion criteria and based on the quality assessment. To identify each research paper in this SLR, a unique ID was given to each study (P1, P2, P3….). In appendix A it is included a table that contains a list of the primary studies used in the SLR with their corresponding ID.

4.1 Overview of studies

This section of the research describes the overview of the 27 primary studies. To have a better understanding of these papers, they were classified based on their research method, type of mobile device used for studying the playability of mobile games, and type of game that was evaluated in the research. Also, the date of publication was considered as part of this analysis.

4.1.1 Publication trend

In this study, the publication year was not used as an inclusions/exclusions criteria. However, it was interesting to observe that the studies had been published in recent years, between 2010 and 2021.

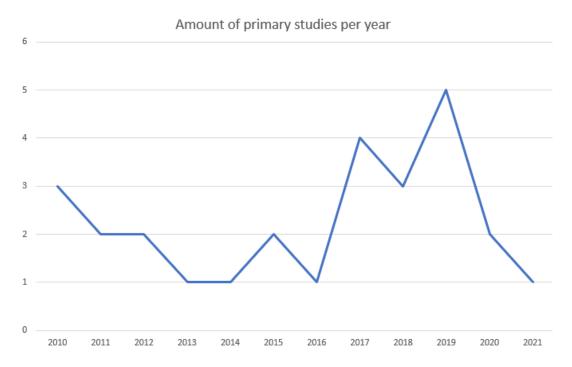**Figure 3.** Distribution of primary studies published by year

As it is observed in Figure 3, there was an increase in the number of studies published from 2016 to 2019. However, there was a decrease in the number of primary studies published in 2021. The reason for this could be due to the fact that the SLR was conducted in April 2021. From the distribution of primary studies published, it can be stated that

there is a positive trend related to the study of usability and playability in mobile games. Therefore, the playability of mobile games can be considered still a relevant topic.

4.1.2 Research Methods

From the 27 primary studies, 7 different research methods were detected, the overview of the results is summarized in Table 6.

Table 6.   Frequency of research methods used in the primary studies

#	Research Method	Number of studies	References
1	Design Science Research	9	P3, P15, P16, P17, P19, P20, P22, P25, P27
2	Experiment	6	P1, P6, P7, P8, P12, P23
3	Literature Review	2	P9, P13
4	Systematic Literature Review	1	P5
5	Survey	3	P2, P4, P24
6	Questionnaire	1	P21
7	Case study	5	P10, P11, P14, P18, P26
Total		27	

The research method that was detected the most, which comprised 32% of the total of the studies, was Design Science Research. The studies that used this research method developed and tested a mobile game based on a literature review. The second most used research method identified, which comprised 25% of the total studies was experimentation. In this type of research, the focus of the study was to do an experiment to evaluate the usability/playability of mobile games. This was followed by case study (18%), where the usability/playability of a game was studied. The less common research methods found from the primary studies were surveys and questionnaires (14%), literature reviews (7%), and systematic literature review (3%).

4.2 Evaluated Game Type

In this section, the primary studies were categorized based on the type of game the paper used for evaluating the playability of mobile games, see Table 7. The primary studies were classified based on the game genre that was proposed by Vince (2018) plus a category for serious game, which was not included in the original list of categories. From the SLR 19 studies were detected where the playability of a mobile game was assessed, and 10 different game genres were found.

Table 7. Primary studies classified based on type of games

Evaluated game type	Number of studies	References
Serious game	9	P13, P16, P17, P18, P19, P22, P25, P26, P27
Racing game	3	P4, P7, P23
Platform	1	P12
Simulation	1	P10
Rhythm game	1	P24
Drawing game	1	P15
Puzzle game	1	P5
Board game	1	P20
Sport game	1	P11

The type of game that was most present in the primary studies was serious game P13, P16, P17, P18, P19, P22, P25, P26, P27, followed by racing games P4, P7, P23. It was interesting to notice that one of the studies used augmented reality (AR) in the game P20. According to that study, the users found the game more attractive to play in comparison with 2D games P20.

4.3 Mobile device type

The primary studies were categorized based on the type of mobile device study: smartphone, tablets, and mobile device. The category mobile device was applied for the primary studies where it was not specified the kind of mobile device that was used in the research or if the research was about mobile devices in general, which includes tablets and smartphones. Table 8 shows the distribution of primary studies based on the type of device that was used for the focus of the research.

Table 8. Primary studies classified based on type of mobile device

Mobile Device Type	Number of papers	References
Smartphone	19	P1, P2, P3, P5, P6, P7, P8, P9, P10, P12, P14, P16, P19, P20, P21, P23, P25, P26, P27
Tablet	3	P11, P18, P22
Mobile device	5	P4, P13, P15, P17, P24

The device that was presented the most in the primary studies was the smartphones which represent 71% of the total primary studies population. This showed that the study of the playability of mobile games is done mainly on smartphones. On the other hand, only 3 studies, which represented 10% of the total of the selected primary studies, used a tablet for researching the playability of mobile devices. There were 5 studies, which represented 17% of the population, that focused on mobile devices in general.

4.4 Analysis of the results

In the next sections, the results of the SLR which answer to RQ1 (what are the methodologies used to measure playability in mobile games) and RQ2 (is there a recommended methodology for evaluating the playability of mobile games) are presented.

4.5 Methodologies for evaluating playability in mobile devices

According to the data that was collected from the SLR, it identified 12 different methodologies that are used to evaluate the usability and playability of mobile games. Table 9 shows the methodologies that were identified and a reference to the primary study.

Table 9. Primary studies classified based on evaluation methodology

Methodology	Number of studies	References
Heuristic Evaluation	11	P2, P3, P4, P6, P7, P8, P10, P13, P14, P24, P25
Playtesting	9	P1, P2, P3, P11, P12, P13, P18, P25, P26
Inspection of interface	1	P9
Thinking aloud	3	P11, P12, P15
Filmed play	4	P11, P12, P23, P25
Observation of gameplay	3	P12, P15, P20
Focus Groups	1	P12
Usability testing	6	P5, P11, P12, P16, P17, P22
Questionnaire/Post-game questionnaire	6	P16, P17, P22, P23, P26, P27
Interview/Post-game interview	5	P12, P18, P19, P20, P21
Expert Evaluation	3	P2, P21, P13
Data logging	1	P12

The data collected from the SLR showed that the most used methodology for evaluating the playability of mobile games was the heuristic evaluation P2, P3, P4, P6, P7, P8, P10, P13, P14, P24, P25. The second most used methodology for understanding and evaluating the playability of mobile games was playtesting P1, P2, P3, P11, P12, P13, P18, P25, P26. Followed by usability testing, questionnaire/post-game questionnaire, interview/post-game interview, filmed play, observation of gameplay, think-aloud, expert evaluation, inspection of interface, data logging, and focus groups.

The following subchapters contain the most significant findings related to the methodologies that are used to evaluate the playability of mobile games. The methodologies that were only mentioned in the primary studies that did not contain valuable information were not included.

4.5.1 Usability Testing

Mozgovoy and Pyshkin (P11) state that the purpose of conducting usability testing is to check that the interaction between the user with the game is effortless. To be more precise the user can learn the rules of the game, navigate through the graphical user interface to find the functions that they are looking for, and the user is able to concentrate on the game itself instead of navigating through confusing interfaces and learning "counter-intuitive" controls P11.

In the usability tests, the user is recorded while playing the game. Usually, two cameras are utilized: one is used to film the face of the user and the second one is used to record the device where the user is playing. At the end of the session, the film is analyzed to study the users' reactions while playing the game P11. One of the benefits of conducting a usability test is that it is possible to get direct feedback from the end-user P5.

4.5.2 Heuristic Evaluation

Heuristic evaluation is the most used methodology for evaluating the usability and playability of mobile games P2, P3, P4, P6, P7, P8, P10, P13, P14, P24, P25. It consists of a group of usability experts that evaluates a mobile game based on a list of usability and playability principles (heuristics) P13.

The Playability Heuristic for Mobile Games proposed by Korhonen and Elina Koivisto (2006) is still relevant in the scientific community, it was mentioned in the following primary studies P2, P3, P5, P7, P8, P13, P14. Also, their heuristics have been used for evaluating the playability of mobile games P2, P7, P8, P10, P17 or for creating more precise heuristics for specific games genre, P24. In one of the studies, a new set of heuristic was suggested based on the PM framework proposed by Sánchez et al (P3) and in the mobile heuristic evaluation presented by Korhonen and Koivisto (2006). Based on a literature review the author of the study proposed the separation of the game elements that are used for evaluating the playability of mobile games in two categories: "On-Screen" and "Off-Screen". On-Screen refers to the elements that are present visually in the game, such as game interface, mechanics, gameplay, and storyline P3. On the other hand, the "Off-Screen" category contains all the aspects that are not part of the game, such as the environmental and social factors that influence the game experience P3.

Furthermore, it was found that the current playability heuristic lacks certain measurements to identify all playability problems. Therefore, it is required a new set of

playability evaluations to support the new features that mobile devices have nowadays, such as touch screens (P5).

The study P2 focused on identifying the playability problems that were not covered by the heuristic evaluation for mobile games proposed by Korhonen and Koivisto (2006). This was done through a literature review and by interviewing university students who considered themselves as regular players. In addition to the heuristics of Korhonen and Koivisto (2006), it was proposed a set of heuristics for mobile games P2. Table 10 contains the list of heuristics that were proposed in the research.

Table 10. List of proposed heuristics for mobile games obtained from (P2)

Gameplay	
1	The player able to save the game anytime
2	Game objectives are moderate (not to easy-nor to difficult)
Usability	
3	Player able to skip movies & images (non-playable)
4	Game allow customization
Mobility	
5	Game can handle interruptions (internal)
6	Player able to pause the game anytime
Multiplayer	
7	Multiplayer sessions can be easily created
8	Game sessions can be saved & restored in loss of connectivity
9	Game supports multiple connectivity medium
10	Game supports multiple ways of communications (voice & text)

According to Chen & Lo (P24), the heuristics proposed by Korhonen and Koivisto (2006) are too general to apply to a specific genre. For example, Chen & Lo proposed their list of heuristics for evaluating playability in rhythm games. The heuristics that Chen & Lo proposed are divided into three categories: functional, structural, and audiovisual (P24). Another example of this phenomenon is the study written by Sarwar Khan et. al about casual mobile games, where a new set of heuristics for evaluating casual games was presented (P5).

One of the disadvantages of using Heuristic Evaluation is the time it takes to perform this kind of study. However, it was found a study (P3) that developed a playability heuristic evaluation system (PHES) in order to automate this process and reduce the amount of time it takes to conduct a heuristic evaluation. The results of the study were positive and

in fact, it is possible to automate the evaluation of playability of mobile games through a web-based system. Nonetheless, this was just a research, and it seems that there is still no available platform for commercial use yet.

4.5.3 Heuristic Evaluation for controllers in mobile devices

One of the elements that affect the playability of mobile games, is the game controls. There were identified studies that focus on how the game controls affect the playability of mobile games P1, P9.

In the primary study P1, it was evaluated how the playability of mobile games is affected by two different types of controls: soft keypad (touch screen) and hard keypad (buttons). This study was conducted through a playtesting session where the participants play a game using a mobile device with a keypad and another one with a touch screen. From this study, it was concluded that the players who use touch screens got higher scores and barely make mistakes.

The study P9 conducted a literature review related to usability heuristics and based on the results it proposed a set of heuristics for evaluating mobile devices that uses a touch screen. The heuristics proposed by the study P9 are included in Appendix C.

4.5.4 Playtesting

Playtesting was the second most encountered methodology for evaluating the playability of mobile games P1, P2, P3, P11, P12, P13, P18, P25, P26. This methodology is common to use for evaluating the playability of video games P2, P3. However, this methodology is only useful when there is a fully functional prototype of the game P2, P3.

One of the drawbacks of this methodology, in comparison with heuristic evaluation, is the time it takes to perform a playtesting session P2, P3. The reason for this is because the user needs to get familiarized with the game before conducting the test P3. After the gaming session, the user goes through a questionnaire or interview about the game and about its experience while playing the game P3.

As it is mentioned by Mozgovoy and Pyshkin (P11) "the goal of playtesting is to seek explicit player opinion about their satisfaction with the game and willingness to play" (p. 6).

4.5.5 Think-aloud

The think-aloud methodology was used for finding usability issues in mobile games P11, P12, P15. This methodology is a simple and cost-effective way to find issues when the user plays the game P11. As was mentioned by Nielsen (2012), among the benefits of using the think-aloud methodology for evaluating the usability of a product is its low cost. The think-aloud sessions conducted in the study P11 were recorded for further analysis.

4.5.6 Expert Evaluation

The expert evaluation consists of an expert or group of experts that evaluates the usability of a product P23. This methodology is mostly used for testing the usability of software

products. Therefore, it cannot be used directly for evaluating the playability of mobile games since it does not take into account the playability elements, such as gameplay, game story, mobility, multiplayer features, among others. The objectives of games, in comparison with utility software, are different. To use expert evaluation methodology in the mobile game context, it must have a list of specific playability heuristics P2.

4.5.7 Post-game interview/ post-game questionnaire

Another methodology that was frequently used for evaluating the playability of mobile games was through post-game interview P19, P20, P23, and post-game questionnaire P17, P18, P20, P22, P23, P27. In this methodology, an interview or questionnaire is given to the participant at the end of a gaming session. The most common way of conducting this methodology according to the extracted data is through a questionnaire. However, there were two studies (P20, P23) that used a mix of post-game questionnaire following by a post-game interview.

The structure of the questionnaire/interview of the study P18 was divided into three sections. In the first section it asks questions related to the participant demographic, the second section contains questions related to the playability/usability of the mobile game, and the last section contains open questions related to the experience of the participant while using the game.

While the objective of most post-game interviews/post-game questionnaires of the primary studies was to assess the playability and usability of mobile games, there was a study that focuses mainly on the usability of the game without taking into account its playability P22.

4.5.8 Filmed play

This methodology consists of recording the test user while playing the mobile game. The player is recorded using two cameras, one camera that is used to record the face of the participant, and the second camera is used to record the screen of the device P11, P25.

It was interesting to observe one study that used the camera of the mobile device to record the facial expressions of the participant P23. However, it was not able to record the screen of the mobile device.

As it was stated in studies P11 and P23, the video records are used to analyze the interaction of the player with the mobile game and possibly detect usability/playability issues.

This methodology can be used with other methodologies simultaneously without interfering, such as was observed in the study P11. This study it was used a combination of usability testing, thinking-aloud, and filmed play to capture the playability and usability issues of their game.

4.6 Recommended methodologies for evaluating the playability of mobile games

The most encountered methodologies for evaluating the playability of mobile games in this SLR was heuristic evaluation P2, P3, P4, P6, P7, P8, P10, P13, P14, P24, P25 and playtesting P1, P2, P3, P11, P12, P13, P17, P18, P25, P26.

From the identified methodologies for evaluating the playability of mobile games, the most effective methodology was the heuristic evaluation. As it was stated by Soomro et al. (P3) "heuristics evaluation is proven to be more efficient to evaluate games with playability heuristics as compared to other methods such as playtesting" P3.

Another aspect that was found was that it is possible, with the use of Korhonen and Koivisto's (2006) heuristics to identify a considerate amount of playability issues of mobile games. Nonetheless, studies that have focused on the development of more specific heuristics for games of certain genres had found that it is not possible to identify all the playability issues by using the current heuristics evaluations. Therefore, they have proposed a new set of heuristics for that specific game genre P4, P13, P24.

Although heuristic evaluation and playtesting are the most effective methodologies for evaluating the playability of mobile games, it is feasible to evaluate this phenomenon using other methodologies, such as expert evaluation. However, it must include playability heuristics as part of the evaluation P2.

5.   Discussion

In this section, it is discussed general observations of this SLR, followed by the answers to the research questions.

A total of 1,390 studies related to the playability of mobile games were found. From which 27 papers were identified as primary studies for this research. These primary studies were classified based on the research method, the type of game, the type of mobile device, and the methodology used for evaluating the playability of mobile games. The most used research method identified in this SLR was Design Science Research, followed by experiment and case study. The most used device for evaluating the playability of mobile games was the smartphone. From the collected data it can be inferred that the smartphone is the most popular mobile device for playing mobile games. Among the reasons that could contribute to this are its mobility and its affordability. The most encountered type of mobile game in this SLR was serious games. This shows that more studies are done to understand how mobile games can be used to also teach new skills and knowledge.

Regarding the RQ1: What are the methodologies that are used to measure the playability of mobile games? There were identified 12 different methodologies, summarized in Table 9. The most encountered methodologies in this SLR were playability heuristic evaluation, playtesting, usability testing, post-game questionnaire/post-game interview, and filmed play. It was interesting to observe the expert evaluation and the inspection of the interface were among the less frequent methodologies in this SLR. The reason for this could be due to the fact that it requires an expert in usability that evaluates the game. Also, another reason why these methodologies were not that common is that they lack the encounter with the final user, which can give valuable feedback related to the usability and playability of the game.

Concerning the RQ2: Is there a recommended methodology for evaluating the playability of mobile games? The methodology that was considered the most appropriate for evaluating the playability of mobile games is the playability heuristic evaluation. The heuristic evaluation for mobile games proposed by Korhonen and Koivisto (2006) can be used for evaluating the playability of most types of games. However, in order to perform a more complete evaluation of the playability of certain types of mobile games, it is recommended to used more specific heuristics.

Another aspect that needs to be taken into account before using the playability heuristics for mobile games is that it does not take into account touch screen devices. The evaluation of the playability of mobile games based on touch screens seems to be an area that still needs to be further researched. Soomro et al. (P8) proposed, "the development of a new set of playability heuristics that support touch screen usability and mobility issues for mobile phones" (p. 273). However, in the research written by Daud et al. (P9) it was identified a proposal of a set of heuristics for evaluating the controls of mobile devices that use touch screen.

One of the main drawbacks of performing a heuristic evaluation is the time it takes to do. However, it was encountered one study (P3) where it was developed a system for automating the heuristic evaluation. This study showed that it is possible to automate the evaluation of playability of mobile games and therefore reduce the time it takes to perform a heuristic evaluation.

Among the reasons why heuristic evaluations and playtesting were considered to be more appropriate, comparing with expert evaluation is that they can get feedback from the end-user. This allows a better understanding of the user experience and usability problems the player encounters while playing the game. Therefore, it is recommended the use of these kinds of heuristics for a more complete evaluation of the playability of the game. Nonetheless, it is possible to perform an expert evaluation for assessing the playability of mobile games. The only requirement is that the evaluation needs to be performed based on a set of playability heuristics.

6.   Conclusion

The purpose of this thesis is to understand how it is evaluated the playability of mobile games. To do this, it was performed an SLR using three databases: Web of Science, Scopus, and IEEE Xplore. From the databases were identified 1,390 studies related to the playability of mobile games, from which 27 papers were recognized as primary studies.

The result of this thesis is the identification of 12 methodologies that are used for evaluating the playability of mobile games. The most used methodologies for understanding the playability of mobile games were heuristic evaluation and playtesting. The playability heuristics for mobile games can be used to evaluate most mobile games. However, to have a more complete evaluation of a game from a specific genre, it is required more specialized heuristics.

The most effective methodologies for evaluating the playability of mobile games are heuristic evaluation and playtesting. Other methodologies, such as expert evaluation, can be used for assessing the playability of mobile games. However, it is required for the methodology to contain a list of heuristics that can be used to study the playability of the game.

The results of this thesis are valuable for game developers, game designers, and game usability practitioners who are interested in understanding the methodologies that are used for evaluating the playability of mobile games. The contribution of this thesis is the summarizing of the current methodologies that are used to understand and evaluate this aspect of game usability.

6.1 Study limitations

This study was conducted using the SLR methodology. The purpose of using this methodology is to minimize the bias in the research (Kitchenham & Charters, 2007). However, here is a list of the limitations of this study:

- The study was performed by a master's degree student who has taken a course related to usability
- The study was conducted using three databases: IEEE Xplore, Scopus, and Web of Science. However, there are other databases that were not included in the study that might have valuable information related to this topic, such as Springer and ACM DL Digital Library
- The primary studies that were not available were ignored
- The study was performed by one person

The threats of this study can be found in the process of the selection of primary studies and the data extraction. It is recommended by Kitchenham & Charters (2007) that the selection of primary studies should be performed by more than one assessor to reduce bias in the study. To minimize this threat, a protocol for conducting the SLR was developed and evaluated by the master thesis supervisor. In the protocol was specified the inclusion and exclusion criteria that were used to select the primary study. Furthermore, to reduce bias during the selection of the primary studies, the title and abstracts of the studies were read several times to make sure the relevant studies are included in the SLR.

To reduce biases during the data extraction, Kitchenham & Charters (2007) suggests the use of data extraction forms to record in detail the information obtained from the primary studies. After the form is defined in the protocol, it should be piloted (Kitchenham & Charters, 2007). Due to the nature of this study, it was considered that an excel spreadsheet should be enough for collecting the data of the primary studies. Another aspect that was not performed according to Kitchenham & Charters guidelines for conducting an SLR was the piloting of the extraction form (excel sheet). This can propose a thread to the quality of this study and can cause some biases in the data extraction process.

6.2 Future research

The focus of this research was to identify the methodologies that are used by the community to evaluate the playability of mobile games (smartphones and tablets). However, the studies that research the playability of mobile games that use VR and wearable devices as controllers were not taken into consideration. Therefore, it would be interesting to conduct a research focused on these devices to understand the limitations of current playability evaluation methodologies

Another aspect that would be recommended for further studies is the research of the effectiveness of different methodologies for evaluating the playability of mobile games. Most of the studies only focus on one methodology. But would be nice a comparative study to evaluate the effectiveness of each methodology in evaluating mobile games and under which circumstances they should be used.

From this SLR it was identified a lack of studies related to the evaluation of playability of mobile games that use touch screens. Although nowadays most mobile games depend on touch screens to control the game, the current heuristics does not take that into account. The understanding of how the touch screen of mobile devices affects the playability and the player experience would be beneficial to the gaming community. In the end, the main goal of producing games is to create a fun and pleasant experience for the player.

7.   References

Wang, A. I., Wu, B. & Bakken., S. K. (2010). Experiences from implementing a face-to-face educational game for iPhone/iPod touch. Paper presented at the - 2010 2nd International IEEE Consumer Electronics Society's Games Innovations Conference, 1-8. doi:10.1109/ICEGIC.2010.5716895

Chua, A. Y. K., Banerjee, S., & Pee, L. G. (2015). An m-leaming game for the study of humanities. Paper presented at the - 2015 IEEE Conference on E-Learning, E-Management and E-Services (IC3e), 137-142. doi:10.1109/IC3e.2015.7403501

Abdulhak, S. A., Cha, K., & Kang, D. (2011). Evaluation of online social network games usability using verbal protocol analysis. Paper presented at the Seogwipo, Jeju, South Korea. 7-11.

Aker, C., Rizvanoglu, K., Inal, Y., & Yilmaz, A. S. (2016). Analyzing playability in multi-platform games: A case study of the fruit ninja game. Design, User Experience, and Usability: Novel User Experiences, Pt Ii, 9747, 229-239. doi:10.1007/978-3-319-40355-7_22

Alhaidary, R., & Altammami, S. (2017). Can the success of mobile games be attributed to following mobile game heuristics? Social Computing and Social Media: Human Behavior, Pt I, 10282, 13-21. doi:10.1007/978-3-319-58559-8_2

Barnett, L., Harvey, C., & Gatzidis, C. (2018). First time user experiences in mobile games: An evaluation of usability. Entertainment Computing, 27, 82-88. doi:10.1016/j.entcom.2018.04.004

Wong, C. Y., Chu, K., Khong, C. W., & Lim, T. Y. (2010). Evaluating playability on haptic user interface for mobile gaming. Paper presented at the - 2010 International Symposium on Information Technology, 3 1093-1098. doi:10.1109/ITSIM.2010.5561513

Cambridge Dictionary.Usability. Retrieved from https://dictionary.cambridge.org/

Clement, J. (2021). Number of games released on steam worldwide from 2004 to 2020. Retrieved from https://www.statista.com/statistics/552623/number-games-released-steam/

Daud, E. H. C., Mokhtar, S. A., & Mohd, F. (2016). Usability heuristics in the context of control features on mobile games

Davis, J. P., Steury, K., & Pagulayan, R. (2005). A survey method for assessing perceptions of a game: The consumer playtest in game design. Game Studies, 5, 1-13.

Dörner, R., Göbel, S., Effelsberg, W., & Wiemeyer, J. (2016). Serious games. Springer.

Fabricatore, C. (2007). Gameplay and game mechanics: A key to quality in videogames. ENLACES (MINEDUC Chile) -OECD Expert Meeting on Videogames and Education, , 29-31.

González Sánchez, J. L., Gutierrez Vela, F. L., Montero Simarro, M., & Padilla-Zea, N. (2012). Playability: Analysing user experience in video games. Behaviour & Information Technology, 31(10), 1033-1054. doi:10.1080/0144929X.2012.710648

González Sánchez, J. L., Padilla Zea, N., & Gutierrez, F. L. (2009). Playability: How to identify the player experience in a video game. Human-Computer Interaction – INTERACT 2009, 356-359.

Guardiola, E. (2019). Gameplay definition: A game design perspective. Cologne Game Lab,

Hemingway, C., Baja, E. S., Dalmacion, G. V., Medina, P. M. B., Guevara, E. G., Sy, T. R., Taegtmeyer, M. (2019). Development of a mobile game to influence behavior determinants of HIV service uptake among key populations in the philippines: User-centered design process. JMIR Serious Games, 7(4) doi:10.2196/13695

J. S. Khan, F. Arif, & A. Mohsin. (2015). Usability concerns of android casual game applications: Analysis and improvements. Paper presented at the - 2015 Science and Information Conference (SAI), 206-216. doi:10.1109/SAI.2015.7237146

Kitchenham, B. (2004). Procedures for performing systematic reviews., 1-28.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report EBSE-2007-01, 1-57.

Korhonen, H. (2011). The explanatory power of playability heuristics

Korhonen, H. (2016). Evaluating playability of mobile games with the expert review method. Dissertations in Interactive Technology, 24 Retrieved from https://trepo.tuni.fi/handle/10024/99584

Korhonen, H., & Koivisto, E. M. (2006). Playability heuristics for mobile games. Paper presented at the Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services, 9-16.

Matshafeni, L. L. (2017). The playability of a selected computer game and its impact towards the improvement of english vocabularies. Paper presented at the Umhlanga. doi:10.1109/ICTAS.2017.7920522 Retrieved from https://ieeexplore-ieee-org.pc124152.oulu.fi:9443/document/7920522

Mozgovoy, M., & Pyshkin, E. (2018). A comprehensive approach to quality assurance in a mobile game project doi:10.1145/3290621.3290835

Mylly, S., Rajanen M., & Iivari, N. (2019). Usable usability heuristics for game developers. Paper presented at the Toulon, France. Retrieved from https://aisel.aisnet.org/isd2014/proceedings2019/NewMedia/4/

Nacke, L. E., Drachen, A., Kuikkaniemi K., Niesenhaus, J., Korhonen, H., van den Hoogen, Wouter M., . . . de Kort, Yvonne A. W. (2009). Playability and player experience research. Paper presented at the Proceedings of Digra 2009: Breaking New Ground: Innovation in Games, Play, Practice and Theory. Retrieved from https://link.springer.com/content/pdf/10.1007/978-3-642-03655-2_39.pdf

Nielsen J. (2012). Thinking aloud: The #1 usability tool. Retrieved from https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/

Nørgaard, M., & Hornbæk, K. (2006). What do usability evaluators do in practice? an explorative study of think-aloud testing. Paper presented at the Proceedings of the 6th Conference on Designing Interactive Systems, 209-218.

Novick, D., Vicario, J., Santaella, B., & Gris, I. (2014). Empirical analysis of playability vs. usability in a computer game. Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments, 8518, 720-731.

Olsen, T., Procci, K., & Bowers, C. (2011). Serious games usability testing: How to ensure proper usability, playability, and effectiveness. Paper presented at the International Conference of Design, User Experience, and Usability, 625-634.

Ong, J. G., Lim-Ashworth, N. S., Ooi, Y. P., Boon, J. S., Ang, R. P., Goh, D. H., . . . Fung, D. S. (2019). An interactive mobile app game to address aggression (regnatales): Pilot quantitative study. JMIR Serious Games, 7(2) doi:10.2196/13242

Paavilainen, J., Alha, K., & Korhonen, H. (2012). Exploring playability of social network games. Paper presented at the Kathmandu, Nepal, 7624 336-351. Retrieved from https://link-springer-com.pc124152.oulu.fi:9443/chapter/10.1007%2F978-3-642-34292-9_24

Pinelle, D., Wong, N., & Stach, T. (2008). Heuristic evaluation for games: Usability principles for video game design. Paper presented at the Florence, Italy. 1453-1562. Retrieved from https://dl-acm-org.pc124152.oulu.fi:9443/doi/10.1145/1357054.1357282

Ponnada, A., & Kannan, A. (2012). In Thampi, SM ElAlfy, ES Aguiar,J. (Ed.), Evaluation of mobile games using playability heuristics

Alifia, R. R., Rahma, F., Hamida, U., Koesoema, A. P. , Irawan, Y. S., & Egawa, S. (2020). Tsunami fighters: Collaborative multilingual mobile game for earthquake and tsunami disaster preparedness education. Paper presented at the - 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC), 1-6. doi:10.1109/R10-HTC49770.2020.9357038

Robson, R. S., & Sabahat, N. (2020). Heuristic based approach for usability evaluation of mobile games. Paper presented at the - 2020 International Conference on Computing, Electronics & Communications Engineering (iCCECE), 156-161. doi:10.1109/iCCECE49321.2020.9231175

Rajanen M., & Marghescu, D. (2006). The impact of game usability to player attitude.

Rajanen M., & Rajanen D.. (2017). Usability benefits in gamification. Paper presented at the GamiFIN, Pori, Finland. 87-95.

Rajanen, M., & Nissinen, J. (2015). Survey of game usability practices in northern european game companies. Information Systems Research Seminar, 8

Rajanen, M., & Rajanen, D. (2017). Usability benefits in gamification. Paper presented at the Pori, Finland. 87-95.

Rebollo, C., Remolar, I., Rossano, V., & Lanzilotti, R. (2021). Multimedia augmented reality game for learning math. Multimedia Tools and Applications, doi:10.1007/s11042-021-10821-3

Riwinoto, & Pertiwi, M. (2019). Playability of game joko run. Paper presented at the Proceedings of the 2019 2nd International Conference on Applied Engineering, doi:10.1109/ICAE47758.2019.9221729

Rocha, F., Pereira, L. L., Santa, P. M., Roque, L., & Cardoso, J. C. S. (2019). Mapping controls on a 2D user drawn racetracks driving game – an usability assessment. Paper presented at the 653-660. doi:10.1145/3341215.3356302

Koceski, S., & Koceska, N. (2011). Interaction between players of mobile phone game with augmented reality (AR) interface. Paper presented at the - 2011 International Conference on User Science and Engineering (i-USEr), 245-250. doi:10.1109/iUSEr.2011.6150574

Kothandapani, S., Goh, D. H., & Razikin, K. (2012). Influence of playability and usability in a mobile human computation game. Paper presented at the - 2012 32nd International Conference on Distributed Computing Systems Workshops, 75-78. doi:10.1109/ICDCSW.2012.44

Soomro, S., Ahmad, W. F. W. & Sulaiman, S. (2014). Evaluation of mobile games with playability heuristic evaluation system. Paper presented at the - 2014 International Conference on Computer and Information Sciences (ICCOINS), 1-6. doi:10.1109/ICCOINS.2014.6868403

Senap, N. M. V., & Ibrahim, R. (2019). A review of heuristics evaluation component for mobile educational games. Fifth Information Systems International Conference, 161, 1028-1035. doi: 10.1016/j.procs.2019.11.213

Sicart, M. (2008). Defining game mechanics. The International Journal of Computer Game Research, 8(2) Retrieved from http://gamestudies.org/0802/articles/sicart

Song, D. H., Kim, K. B., & Lee, J. H. (2019). Analysis and evaluation of mobile rhythm games: Game structure and playability. International Journal of Electrical and Computer Engineering, 9(6), 5263-5269. doi:10.11591/ijece.v9i6.pp5263-5269

Soomro, S., Ahmad, W. F. W., & Sulaiman, S. (2013). Evaluation of mobile games using playability heuristics. Advances in Visual Informatics, 8237, 264-274.

Soomro, S., Wan Ahmad, W. F., & Sulaiman, S. (2013). Evaluation of mobile games using playability heuristics. Advances in Visual Informatics. IVIC 2013. Lecture Notes in Computer Science, 8237, 264-274.

Szklanny, K., Homoncik, Ł., Wichrowski, M., & Wieczorkowska, A. (2017, September). Creating an interactive and storytelling educational physics app for mobile devices. In 2017 Federated Conference on Computer Science and Information Systems (FedCSIS) (pp. 1269-1273). IEEE.

testbytes. (2019). What is PlayTesting? how to get good feedback for your game? Retrieved from https://www.testbytes.net/blog/what-is-play-testing/

Uggah, L. L., Manaf, A. A. A., & Adis, A. -. A. (2018). Evaluating south korean based mobile role playing games with playability heuristic evaluation. International Journal of Engineering and Technology (UAE), 7(3), 1-3. doi:10.14419/ijet.v7i3.18.16661

Senap, N. M. V., & Ibrahim, R. (2019). A review of heuristics evaluation component for mobile educational games. Procedia Computer Science, 161, 1028-1035.

Vince. (2018, April 12,). The many different types of video games & their subgenres. Retrieved from https://www.idtech.com/blog/different-types-of-video-game-genres

Wijman, T. (2020). The world's 2.7 billion gamers will spend $159.3 billion on games in 2020; the market will surpass $200 billion by 2023. Retrieved from https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023/

Chee, Y. S., Tan, E. M., & Liu, Q. (2010). Statecraft X: Enacting citizenship education using a mobile learning game played on apple iPhones. Paper presented at the - 2010 6th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technologies in Education, 222-224. doi:10.1109/WMUTE.2010.16

Xu, Z., & Qiu, G. (2017). Draw-and-guess: A mobile game with a purpose for drawing base image retrieval. Paper presented at the - 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 202-207. doi:10.1109/ICMEW.2017.8026305

Appendix A. List of Primary Studies

ID	Primary Study	Authors
P1	Evaluating playability on haptic user interface for mobile gaming	Wong, C. Y., Chu, K., Khong, C. W., & Lim, T. Y.
P2	A preliminary study on heuristics for mobile games	Soomro, S., Ahmad, W. F. W., & Sulaiman, S.
P3	Evaluation of mobile games with playability heuristic evaluation system	Soomro, S., Ahmad, W. F. W., & Sulaiman, S.
P4	Heuristic Based Approach for Usability Evaluation of Mobile Games	Robson, R. S. & Sabahat, N.
P5	Usability concerns of android casual game applications: Analysis and improvements	Khan, J. S., Arif, F., & Mohsin, A.
P6	The Explanatory Power of Playability Heuristics	Korhonen H.
P7	Evaluation of Mobile Games using Playability Heuristics	Ponnada, A., Kannan, A.
P8	Evaluation of Mobile Games Using Playability Heuristics	Soomro, S., Ahmad, W.F.W., & Sulaiman, S.
P9	Usability Heuristics in the Context of Control Features on Mobile Games	Daud, E.H.C., Mokhtar, S.A., & Mohd, F.
P10	Can the Success of Mobile Games Be Attributed to Following Mobile Game Heuristics?	Alhaidary, R., & Altammami, S.
P11	A Comprehensive Approach to Quality Assurance in a Mobile Game Project	Mozgovoy M., & Pyshkin, E.
P12	First Time User Experiences in mobile games: An evaluation of usability	Barnett, L., Harvey, C., & Gatzidis, C.
P13	A Review of Heuristics Evaluation Component for Mobile Educational Games	Senap, N.M.V, & Ibrahim, R.
P14	Evaluating South Korean based mobile role playing games with playability heuristic evaluation	Uggah, L. L., Manaf, A. A. A., & Adis, A. A. A.
P15	Draw-and-Guess: A mobile game with a purpose for drawing base image retrieval	Xu, Z., & Qiu, G.

P16	Experiences from implementing a face-to-face educational game for iPhone/iPod Touch	Wang, A. I., Wu, B., & Bakken, S. K.
P17	Tsunami Fighters: Collaborative Multilingual Mobile Game for Earthquake and Tsunami Disaster Preparedness Education	Alifia, R. R., Rahma, F., Hamida, U., Koesoema, A. P., Irawan, Y. S., & Egawa, S.
P18	An m-learning game for the study of humanities	Chua, A. Y. K., Banerjee, S., & Pee, L. G.
P19	Statecraft X: Enacting Citizenship Education Using a Mobile Learning Game Played on Apple iPhones	Chee, Y. S., Tan, E. M. & Liu, Q.
P20	Interaction between players of mobile phone game with augmented reality (AR) interface	Koceski, S., & Koceska, N.
P21	Usability Evaluation Methods of User Interface Based on Mobile Games Using Fuzzy Methods	Cui M., & Zhu, L.
P22	Creating an interactive and storytelling educational physics app for mobile devices	Szklanny, K., Homoncik, L., Wichrowski, M. & Wieczorkowska, A.
P23	Mapping controls on a 2D user drawn racetracks driving game – An usability assessment	Rocha, F., Pereira, L. L., Santa, P. M., Roque, L., & Cardoso, J. C. S.
P24	Analysis and evaluation of mobile rhythm games: Game structure and playability	Song, D. H., Kim, K. B., & Lee, J. H.
P25	Development of a mobile game to influence behavior determinants of HIV service uptake among key populations in the philippines: User-centered design process	Hemingway, C., Baja, E. S., Dalmacion, G. V., Medina, P. M. B., Guevara, E. G., Sy, T. R., Dacombe, R., Dormann, C., & Taegtmeyer, M.
P26	An interactive mobile app game to address aggression (regnatales): Pilot quantitative study	Ong, J. G., Lim-Ashworth, N. S., Ooi, Y. P., Boon, J. S., Ang, R. P., Goh, D. H., Ong, S. H., & Fung, D. S.
P27	Multimedia augmented reality game for learning math	Rebollo, C., Remolar, I., Rossano, V., & Lanzilotti, R.

Appendix B Heuristic for evaluating mobile games

Game Usability Heuristics	
No.	*Heuristic Evaluation*
GU1	Audio-visual representation supports the game
GU2	Screen layout is efficient and visually pleasing
GU3	Device UI and game UI are used for their own purposes
GU4	Indicators are visible
GU5	The player understands the terminology
GU6	Navigation is consistent, logical, and minimalist
GU7	Control keys are consistent and follow standard conventions
GU8	Game controls are convenient and flexible
GU9	The game gives feedback on the player's actions
GU10	The player cannot make irreversible errors
GU11	The player does not have to memorize things unnecessarily
GU12	The game contains help
Mobility Heuristics	
No.	*Heuristic Evaluation*
MO1	The game and play sessions can be started quickly
MO2	The game accommodates with the surroundings
MO3	Interruptions are handled reasonably
Heuristics for evaluating gameplay	
No.	*Heuristic Evaluation*
GP1	The game provides clear goals or supports player-created goals
GP2	The player sees the progress in the game and can compare the results
GP3	The players are rewarded, and rewards are meaningful
GP4	The player is in control

GP5	Challenge, strategy, and pace are in balance
GP6	The first-time experience is encouraging
GP7	The game story supports the gameplay and is meaningful
GP8	There are no repetitive or boring tasks
GP9	The players can express themselves
GP10	The game supports different playing styles
GP11	The game does not stagnate
GP12	The game is consistent
GP13	The game uses orthogonal unit differentiation4
GP14	The player does not lose any hard-won possessions

Appendix C Preliminary Model for Usability Heuristics

Controls	
No.	**Heuristic Evaluation**
C1	Recognition rather than recall
C2	Flexibility and efficiency of use
C3	User control and freedom
C4	Error prevention
C5	Visibility of system status
C6	Defined carefully clue to the touch screen and its small in size
C7	Have small number of buttons
C8	Use other alternatives: swipe, tilt as long it is efficient and relevant to the theme and game's design