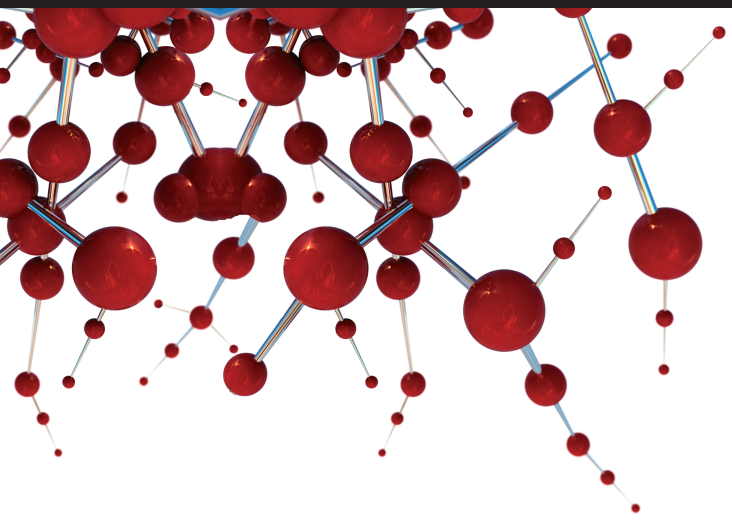
Interactive Technology • School of Information Sciences • University of Tampere



Hannu Korhonen

Evaluating Playability of Mobile Games with the Expert Review Method

Hannu Korhonen

Evaluating Playability of Mobile Games with the Expert Review Method

ACADEMIC DISSERTATION
To be presented with the permission of the School of Information Sciences of the University of Tampere, for public discussion in the Pinni auditorium B1096 on August 26th, 2016, at noon.

School of Information Sciences
University of Tampere

Dissertations in Interactive Technology, Number 24
Tampere 2016

ACADEMIC DISSERTATION IN INTERACTIVE TECHNOLOGY

Supervisors:	Professor Kari-Jouko Räihä, School of Information Sciences, University of Tampere, Finland Professor Frans Mäyrä, Information Studies and Interactive Media, School of Information Sciences, University of Tampere, Finland
Opponent:	Regina Bernhaupt, Ph.D. Associated Researcher in Interactive Critical Systems, Toulouse Institute of Computer Science Research, University of Toulouse III Paul Sabatier, France
Reviewers:	Effie Lai-Chong Law, Ph.D. Reader in Human-Computer Interaction, Department of Computer Science, University of Leicester, The United Kindom Magy Seif El-Nasr, Ph.D. Associate Professor in the Colleges of Computer and Information Sciences and Arts, Media and Design, Northeastern University, The United States of America

The originality of this thesis has been checked using the Turnitin Originality Check service in accordance with the quality management system of the University of Tampere.

Dissertations in Interactive Technology, Number 24

School of Information Sciences
FIN-33014 University of Tampere
FINLAND

ISBN 978-952-03-0196-5
ISSN 1795-9489

Juvenes Print – Suomen Yliopistopaino Oy
Tampere 2016

Acta Electronica Universitatis Tamperensis 1701
ISBN 978-952-03-0205-4 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Abstract

The success of a video game is measured based on the experiences it provides for the players. The players continue playing the game if they feel that it is an enjoyable experience. Usually there are several video games with a similar concept on the market, and players favor those games which provide the most interesting and engaging gameplay. Playability is another aspect which enhances a positive player experience. The player should be able to interact with the gameworld and possibly with other players in a smooth and unobtrusive manner.

Making a good game is a long process and game designers should evaluate both player experience and playability aspects regularly. Player experience is difficult to assess until the players are able to play a balanced and playable game. Playability evaluations can be done throughout the process, but are often mixed up with quality assurance testing even though they have a different focus. In addition, analytical inspection methods, which do not include end-users, are not commonly used in game projects, although they are efficient evaluation methods.

This dissertation addresses issues with an analytical inspection method called expert review when it is used in playability evaluations of a game. The focus of the research work is twofold and it is located at the intersection of evaluation methodology in human-computer interaction and game design. First, the objective is to explore what playability means in video games. The second objective is to define playability heuristics that could be used with the expert review method to evaluate video games during the design and implementation phases of a game development project.

The work consists of analytical and experimental parts. Game design literature was studied to find out the most important aspects of the design that influence the player's interaction with the game and should be included in the playability heuristics set. The experimental part consists of several game evaluations in which the heuristics were used together with the expert review method to evaluate video games. The game evaluations were focused on mobile games as they were Nokia's interest in 2005-2008 when the research work was ongoing. However, the heuristics are applicable for evaluating video games on different gaming platforms.

The results indicate that expert review is a viable method and playability heuristics help inspectors to discover problems in design and thus, improve the playability of a video game.

Acknowledgements

The most important support for my research work has come from my closest colleague Elina Ollila during the first years of the Ph.D. studies. Your knowledge and experience of game design helped us a lot in the beginning of the project. Research work needs comments and critics in order to develop. Janne Paavilainen was the right person for this job and helped me to stay focused with the heuristics. Janne has also frequently provided feedback on how the playability heuristics work in different research projects and evaluation assignments from the industry. I would like to thank Janne and Hannamari Saarenpää for being co-authors in some publications for the dissertation.

Working in a research institute and doing academic research for the university can sometimes be a challenge from the supervision point of view. My supervisors Kari-Jouko Räihä and Frans Mäyrä have been understanding and patient with my research and provided help when it was needed. I would also like to thank Prof. Kaisa Väänänen, Dr. Matt Jones, Dr. Juha Arrasvuori and Dr. Juha Lehikoinen for their support when it was not possible to meet the nominated supervisors. I thank the reviewers of the publications for their valuable comments.

Writing a dissertation at the same time with normal work duties requires understanding and support from your employer. Nokia Research Center was an excellent place to do research work. I thank Jyri Huopaniemi, head of the NRC Tampere research lab, and my former research team leaders, Mika Röykkee, Antti Aaltonen, Jaakko Lehikoinen, Tero Jokela, and Viljakaisa Aaltonen for your support to arrange time and money to do research for this dissertation along with our normal research projects.

Finally, I thank the people who have a special place in my heart. First, my loving wife Kirsi. Without your support this dissertation would never have been finished. Your patience and understanding over the years when I have written articles during the nights and been away from home is something that cannot be thanked too much. I love you! In addition, I thank Sonja's and Joona's grandparents Mirja, Matti, Eeva-Liisa, and Aulis for the support and help that they have provided to our family and Kirsi, when I have attended conferences year after year. Finally, Sonja and Joona, you have had a special role in showing that there is life beyond the laptop and academic publications that is important to all of us.

Kangasala, January 30th, 2016
Hannu Korhonen

Preface

This dissertation work originally started from a normal project request that the Nokia Research Center receives from Nokia's business units. At that time, the entertainment business unit was producing games for the N-Gage platform that 3rd party developers developed. Product managers at the business unit wanted to have a method to evaluate the quality of game design and whether a game was ready to be published as an N-Gage game. The request was received in 2005.

At that time, there were game researchers at the Nokia Research Center who had studied games for a long time and developed their own games as well. However, they did not have experience in evaluating games. Therefore, the request from the business unit was transferred to the Usability Team where I was working. My team manager asked who would like to participate in this joint project with the game researchers. I have played video games since 1983 when Commodore 64 was launched and I was also interested in usability evaluation methods. I decided to take this opportunity and volunteer for the project. I did not think that this would turn out to be my dissertation topic, since I was also involved in another project about voice interfaces and I had already studied them for some time.

The project started together with Elina Ollila (former Koivisto) and other game researchers at the Nokia Research Center. As a first task of the project we had to decide which evaluation method would be the best option to fulfill the request of the business unit. The expert review method was an obvious choice because the business unit would not have a team of usability experts or other professionals that would conduct game evaluations with very complicated methods. Neither did they have facilities for a usability laboratory or similar which would enable efficient user testing of games (or playtesting). They needed a simple method that could be easily taught to game producers or other persons and conducting game evaluations should be fast and cost-efficient.

We developed playability heuristics by reviewing game design literature and conducted game evaluations using the expert review method and the playability heuristic set. There were several prototype games available that we evaluated and made recommendations on how they could be improved from a playability perspective. The project was completed in 2007 and we delivered the playability heuristic set and the description of the method to the business unit persons. Our deliverables were put into operation immediately.

After the project I thought that developing playability heuristics and the expert review method was such an interesting task that it would be a shame to leave it behind. I asked my managers whether it would be possible to continue this project as an extra project and expand it to meet the requirements of an academic dissertation. The project where I studied voice user interfaces was also transferred to a business unit in the previous year and the research part had finished earlier than expected. I decided to take playability heuristics as a new topic for my dissertation. I got the permission to continue this game related research and deepen it with a few additional studies. The length of the project is probably longer than a typical dissertation work due to the nature of how the project was done during these years. Working mainly for a dissertation with close colleagues would have been an ideal way of completing the project earlier, but in my case it was not possible after we finished the project for the business unit. Conducting studies and writing articles on top of the normal research projects at work has been time-consuming.

I conducted more experiments to study the expert review method and whether the results are comparable to playtesting, how different playability heuristics work with the method, and deepen my understanding of the playability of mobile games from a game design perspective, and thus creating a contribution both on a conceptual and an empirical level. This dissertation is an account of the work that I have done with the playability of mobile games and how a project of responding to an industry need was expanded to an academic project towards a Ph.D. degree.

Contents

1 INTRODUCTION 1
- 1.1 Motivation 1
- 1.2 Previous Work 4
  - 1.2.1 Development of Playability Heuristics 6
  - 1.2.2 Game Qualities based on Game Critiques 8
  - 1.2.3 Heuristic Evaluation Method 10
- 1.3 Playing Games on Mobile Devices 12
- 1.4 Mobile Games on Feature Phones and Smartphones 13
- 1.5 Objective and Research Questions 15
- 1.6 Contributions 16
- 1.7 Method Review 17
- 1.8 Structure of the Dissertation 18

2 PLAYABILITY OF (MOBILE) VIDEO GAMES 19
- 2.1 Playability ≠ Usability 21
- 2.2 Playability Influences Player Experience 24
- 2.3 Meaningful Play 27
- 2.4 The Elements of a Game 29
  - 2.4.1 Mechanics 30
  - 2.4.2 Game Story and Avatar 31
  - 2.4.3 Aesthetics 32
  - 2.4.4 Game Technology 32
  - 2.4.5 The Interaction between Game Elements 33
- 2.5 Components Forming the Playability of a Game 34
  - 2.5.1 The Gaming Platform 35
  - 2.5.2 The User Interface 36
  - 2.5.3 Gameplay 37

3 EVALUATION METHODS 39
- 3.1 Usability Inspection Methods 39
  - 3.1.1 The Benefits of Usability Inspection Methods 40
  - 3.1.2 The Challenges of Usability Inspection Methods 40
- 3.2 Expert Review 41
- 3.3 A Procedure for Conducting an Expert Review for Games 42
  - 3.3.1 A Game Evaluation Lasts for Three Rounds 43
  - 3.3.2 Task Scenarios Are Seldom Used in Game Evaluations 44
  - 3.3.3 Identifying Playability Problems 45
  - 3.3.4 Conducting the game evaluation 45
  - 3.3.5 A Secretary would Help in Game Evaluations 47
  - 3.3.6 The Duration of the Evaluation Session 48
  - 3.3.7 The Consolidated List of Playability Problems 49
- 3.4 Inspectors 51
- 3.5 Playtesting in Game Development 54

4 HEURISTICS TO AID INSPECTION 57
- 4.1 General Usability Heuristics 57
- 4.2 Domain-Specific Heuristics for different domains 59

4.3 Domain-Specific Heuristics for Video games 61

5 PLAYABILITY HEURISTICS FOR MOBILE GAMES 67

5.1 Game Usability Heuristics 69
5.2 Gameplay Heuristics 87
5.3 Multi-player Heuristics 106
5.4 Mobility Heuristic 114
5.5 Context-Aware Heuristics 119

6 STUDIES 125

6.1 Study I: Playability Heuristics for Mobile Games 126
6.2 Study II: Playability Heuristics for Mobile Multi-player Games 127
6.3 Study III: Playability Heuristics for Context-Aware Mobile Pervasive Games 128
6.4 Study IV: Comparison of Playability Heuristic Sets with Expert Inspectors 129
6.5 Study V: Comparison of Playtesting and Expert Review in a Mobile Game Evaluation 131
6.6 Study VI: Comparison of Playability Heuristics with Novice Inspectors 133

7 CONCLUSIONS 137

7.1 Reviewing the Contributions 138
7.2 The Limitations of the Study 140
7.3 Implications for Practitioners 141
7.4 Final Remarks 142

8 REFERENCES 145

APPENDIX A: 281

List of Publications

This dissertation consists of a summary and the following original publications, reproduced here by permission.

I. Korhonen, H. & Koivisto, E. M. I. (2006). Playability Heuristics for Mobile Games. In *Proceedings of International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'06, Helsinki, Finland)*, 9-16. New York, NY, USA: ACM. doi: 10.1145/1152215.1152218 171

II. Korhonen, H. & Koivisto, E. M. I. (2007). Playability Heuristics for Mobile Multi-player Games. In *Proceedings of the International Conference on Digital Interactive Media in Entertainment and Arts (DIMEA 2007, Perth, Australia)*, 28–35. New York, NY, USA: ACM. doi: 10.1145/1306813.1306828 181

III. Korhonen, H., Saarenpää, H. & Paavilainen, J. (2008). Pervasive Mobile Games – A New Mindset for Players and Developers. In Markopoulos, P., de Ruyter, B., Ijsselsteijn, W., & Rowland, D. (Eds.), *Proceedings of the International Conference on Fun and Games (Fun'n'Games 2008, Eindhoven, The Netherlands),* LNCS 5294, 21-32. Berlin Heidelberg, Germany: Springer-Verlag. doi: 10.1007/978-3-540-88322-7_3 191

IV. Paavilainen, J., Korhonen, H. & Saarenpää, H. (2012). Comparing Two Playability Heuristic Sets with Expert Review Method: A Case Study of Mobile Game Evaluation. Chapter 2 in Lugmayr, A., Franssila, H., Näränen, P., Sotamaa, O., Vanhala, J., & Yu Z., (Eds.) *Media in the Ubiquitous Era: Ambient, Social, and Gaming Media*, 29-52. Hershey, PA, USA: IGI-Global. doi: 10.4018/978-1-60960-774-6.ch002 205

V. Korhonen, H. (2010). Comparison of Playtesting and Expert Review Methods in Mobile Game Evaluation. In *Proceedings of the International Conference on Fun and Games (Fun'n'Games 2010, Leuven, Belgium),* 18-27. New York, NY, USA: ACM. doi: 10.1145/1823818.1823820 233

VI. Korhonen, H. (2011). The Explanatory Power of Playability Heuristics. In *Proceedings of the Advances in Computer Entertainment Technology (ACE'11, Lisbon, Portugal)*, Article No. 40. New York, NY, USA: ACM, doi: 10.1145/2071423.2071473 245

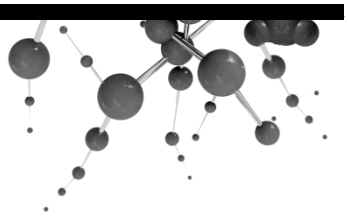
VII. Korhonen, H. (forthcoming). Playability Heuristic Set Influences Inspectors' Performance in a Mobile Game Evaluation, In review for publication in *Entertainment Computing Journal* 255

Author's Contribution to the Research and Publications

The playability heuristics presented in articles I and II were defined together with Elina Ollila. She had the main responsibility of defining Gameplay heuristics and my responsibility was Game Usability heuristics. However, both researchers influenced both heuristic sets. Mobility heuristics and Multi-player heuristics were done in collaboration with Elina with equal contribution. Context-aware heuristics presented in article III were defined by the author of the dissertation. In the evaluation and method validation studies the author has had the main responsibility for the planning of the studies, collecting and analyzing the research data. The only exception is that the study described in article IV was planned in collaboration with Janne Paavilainen. Data collection and analysis were done together with him. In other studies, there were persons helping with arrangements and these persons have been acknowledged in the articles, if they are not co-authors.

The publications included in this dissertation have been written both in collaboration with other researchers and individually by the author. Articles I, II, III and IV were written in collaboration with other researchers. The author of the dissertation had the main responsibility in the writing process for all articles except article IV, which is an extended version of a previously published article (Korhonen, Paavilainen, & Saarenpää, 2009). The article was written in collaboration with Janne Paavilainen and Hannamari Saarenpää. The author's contribution in the writing process is equal with Janne Paavilainen. Hannamari Saarenpää was a third author with a smaller contribution. In the original article, which was published as a conference paper, the author of the dissertation was the primary author. In the extended version of the article, the primary author was switched to Janne Paavilainen. He had the main responsibility in the writing process of the extended parts of the article.

Although other articles V, VI, and VII do not have co-authors, it does not mean that they were written in isolation. Many persons have commented on the draft versions of the articles and they have been acknowledged in the articles.



1 Introduction

1.1 MOTIVATION

Using feature phones and especially smartphones for playing video games has become a common form of entertainment that attracts millions of mobile phone users worldwide. Technological enablers such as increased processing power, graphical accelerators, high resolution displays, and broadband and Wi-Fi connections have enabled us to use mobile phones for playing games. Developing games for mobile devices can be seen as being more popular than game design for game consoles and personal computers (Kersey, 2013).

For every game designer, the main objective is to design a game that is enjoyable and fun for a large player population. However, the design task is not easy and sometimes it takes years to finish the development project of a game. Large development projects are very expensive and time-consuming. Despite all the money and effort that game companies invest in a game, it is not guaranteed that players will like the game.

When we talk about aspects such as enjoyable games that are fun to play, we are typically referring to experiences and how the players perceive the game or any other product (Preece, Rogers, & Sharp, 2007). In the game industry, positive game experience is a crucial factor in determining successful games on the market. There are often several game titles available which are based on similar game concepts. Video games are also one of the few product categories which are constantly reviewed by game critiques, and the players can trial demos before making a purchasing decision. Furthermore, word of mouth is used frequently to receive recommendations of good games from fellow players. If the game experience is not on a satisfactory level, players can easily switch to

another game which hopefully provides a more enjoyable and fun gaming experience.

The players' perception of a game and interaction with the game mechanics will depend on many factors in the design. A player's previous experiences, skills and knowledge of games together with the current context in which the game is played influences the gaming experience. As Donald Norman (2005) has noted, it is not possible to design an experience, but it is possible to design *for* an experience.

Patrick Jordan (2000) has noted that for experience rich products, such as video games, three aspects must be fulfilled to provide a positive user experience: 1) a product should have the right functionality, 2) it must be usable, and 3) it should provide emotional responses to the users. In the video game context, the right functionality could mean that the game concept is interesting to the players. The game is usable when the players do not have to struggle with the game interface, and the game mechanic is understandable to the player. Emotionally stimulating games are those in which the player can feel different emotions such as desire, fear, hope, anger, challenge, relief, anxiety, pride and many more (Fridja, 1986). There is a close relationship between these aspects, and they can be thought of as layers in the design (Jordan, 2000). Without the proper functionality a product will not be usable. Furthermore, a product that is not usable is unlikely to provide desirable emotional responses for the users. In this dissertation, I will focus on usability, or preferably a playability layer for creating a positive user experience, and the issues which make a game playable. Playability is a multifaceted structure which deals with the user interface and the content of the game. Playability is defined and discussed in detail in Section 2.

Designing products that have the right functionality and are usable requires an understanding of the end-users, their tasks or needs, and how they typically interact with the products. This can be done by involving users from the target user group in the design process. This approach is called user-centered design in Human-Computer Interaction literature (ISO 9241-210, 2010). It means that, in every phase of the process, intended users of the product are involved and they can give their input for the product design. Although end-user feedback is valuable throughout the design process, it is especially useful in evaluation tasks in each phase.

A game development project is not an exception in this respect and game designers have collected players' feedback when they evaluate video games. Focus groups are used to evaluate concepts and collect ideas in the concept creation phase of the development project (Fullerton, Swain, & Hoffman, 2004). Playtesting is used to discover difficulties that the players have when playing a game. The players can also provide feedback about

the game experience (Rouse, 2001). Typically, players are invited to playtesting sessions, but mainly playtesting is conducted internally and game designers evaluate partial prototypes by playing them by themselves (Rouse, 2001).

However, there are challenges in this approach. Playtesting by external players cannot be utilized effectively until there is a working prototype available. The prototype usually needs to have a feature-complete status before external players are allowed to see it. Game designers are not very eager to show unfinished games to someone they do not know well (Fullerton, Swain, & Hoffman, 2008, p. 251). The problem arises if serious playability problems are found in a game when it is evaluated by external players. Correcting these problems may require major changes to the design that is usually time-consuming and expensive to do. Leaving them uncorrected will most probably have a negative impact on the game experience, which is a major risk for the success of the game.

Human-Computer Interaction researchers have developed Usability Inspection Methods (UIM) to overcome the problem of discovering serious usability problems too late during product development (Nielsen, 1994b). The methods can be used to evaluate the design when end-user involvement is not yet feasible and provide useful feedback for the designers to redesign problematic features in the product. In usability inspections, the evaluation is conducted by a group of usability experts instead of intended end-users of a product.

Expert review (Nielsen, 1993) is probably the most common inspection method used to evaluate software products. Despite its popularity in software development projects, it has not gained the same esteem in game development projects, or at least the method is not mentioned in game design literature (e.g. Rouse, 2001; Mulligan & Patrovsky, 2003; Fullerton et al., 2004; Fullerton et al., 2008; Schell, 2008) as a viable option for conducting evaluations. Later on, expert review has appeared in the literature as an evaluation method for video games (Isbister & Schaffer, 2008; Shneiderman & Plaisant, 2009).

This dissertation is a multidisciplinary work combining human-computer interaction (HCI) research and game research. HCI research is a broad area and studying users' interaction and performance with a system can be done from many different perspectives. Evaluation methodology research is one research branch, which has been very active since the 1990s and new methods have been defined for different purposes. In this dissertation, instead of defining a new evaluation method, I want to study how we can apply (or possibly modify) the existing method, namely expert review, to suit a new domain.

Game research and game design research define the domain for the dissertation. In order to utilize expert reviews effectively in game evaluations, we need to have domain-specific heuristics to guide the inspection. To define these heuristics, it is essential to understand the characteristics of games. In this dissertation, I will explore existing game research by different game researchers and game design literature to find out the most important characteristics that game designers consider relevant when they are designing games. This will help in determining what kinds of issues domain-specific heuristics should cover.

My approach in the dissertation, especially to game research, is pragmatic. Instead of diving into game research very deeply and getting involved in sometimes heated discussions in the field, I will approach game research from the perspective that practitioners in the game industry would do, to utilize the existing knowledge of research in their daily work. I will explore what aspects constitute the playability of video games and how this knowledge could be used to enhance the expert review method to suit game evaluations better. My motivation for the study is also to promote the expert review method to the game industry as a cost efficient method to evaluate the playability of video games.

1.2 Previous Work

The expert review method has been used to evaluate software for years. Nielsen and Molich (1990) developed the method to lower evaluation costs. The method is cost-efficient and effective in evaluating various types of designs. Prototypes can be partially implemented or even non-functional. The method is commonly used to collect feedback for the designers prior to user testing.

One evaluation round can be completed in a couple of hours, which enables fast and continuous improvement of the design. This is a remarkable saving compared to the user testing procedure, which would require participant recruitment, preparing testing facilities and materials, scheduling tests and conducting tests with participants (Rubin, 1994). User testing will usually take days rather than hours to complete. In addition, studies indicate that a skilful and knowledgeable usability expert can identify usability problems with a similar accuracy as a user testing would provide (Molich & Dumas, 2008). Therefore, it is a tempting idea to evaluate games with this method, but unfortunately it cannot be directly applied to game evaluations.

Nielsen (1994b, p. 29) has indicated that domain-specific heuristics are needed for specific classes of products, such as video games, to supplement general heuristics. Johnson and Wiles (2003) showed in their study how the traditional usability heuristics conflict in game evaluations.

Game researchers (Federoff, 2002; Johnson & Wiles, 2003; Desurvire, Caplan, & Toth, 2004; Korhonen & Koivisto, 2006) have also noted that traditional usability heuristics cover only some of the user interface issues relevant to video games and ignore gameplay aspects completely. Hence, there is a need for developing domain-specific heuristics for game evaluations.

Despite the recognized lack of domain-specific heuristics for video games, game researchers have not studied them extensively. Even a commonly agreed upon definition for playability is missing. The definition should describe high level objectives and guide the research work of developing playability heuristics.

Sometimes playability is referred to as game usability (Isbister & Schaffer, 2008), which would implicate that playability has similar characteristics as usability has in productivity software. According to ISO standard (ISO 9241-11, 1998), usability is defined as effectiveness, efficiency, and user satisfaction in a specified context of use. Effectiveness and efficiency refer to a user's ability and performance in completing tasks with the product and user satisfaction is a result of succeeding in task completion. A usable product is often easy to learn and easy to use (Nielsen, 1993).

Games, however, are complex products and there are more things that influence the playability of a game. The effectiveness and efficiency of the players are not the only sources of user satisfaction and there are usually other design goals than effectiveness and efficiency that game designers want to promote in a game. Enjoyment and fun are achieved in games when they are sufficiently challenging for a player. Learning necessary skills in the game, solving problems, or discovering the gameworld are common instances of challenge in games. Another typical design goal is competition, which includes predefined obstacles and difficulties for players, and unpredictable things to happen before the winner of the game is known. These two design goals are related to the gameworld and they obviously conflict with objectives of designing usable products. Design goals can also relate to astonishing graphics and story which will immerse the players.

On the other hand, effectiveness and efficiency are acceptable objectives for user interface design of the game. In fact, players value user interface design even more if the design is not just efficient, but it is reliable, convenient, and unobtrusive, so that playing the game is as smooth as possible. The player should be able to concentrate on playing the game and not struggle with the user interface. Unfortunately, there have sometimes been incorrect assumptions about what fun means in games. Clanton (1998) says that problems in the game interface and game mechanics can be masked by the fun of solving problems in the game. I would argue that the claim is completely wrong because design problems

in the game interface mean that there are simply basic usability problems and solving them is not a part of a positive gaming experience. Similarly, problems in game mechanics are frustrating to a player, because the player has a mental model and expectations of how the world should operate. If there are problems in the game mechanics those will similarly deteriorate the game experience.

Before domain-specific heuristics can be defined properly, we need to understand what playability means in video games. There are a few earlier studies that have tried to define playability. Egenfeldt-Nielsen, Smith, and Pajares Tosca (2008) have defined that good playability refers to fun, challenge and ease of use. Järvinen, Heliö, and Mäyrä (2002) have defined four components of playability: 1) functional, 2) structural, 3) audio-visual, and 4) social. They can be used as an evaluation tool to evaluate both the formal (game functionalities) and informal (user experience and user practices) aspects of the gameplay and how they interact (Järvinen et al., 2002). Fabricatore, Nussbaum and Rosas (2002) have studied action games and defined playability as the possibility of understanding and controlling the gameplay. According to them, non-functional aspects of the design cannot balance or replace poor playability. Usability Glossary (2002) defines that playability is affected by storyline, controls, pace, usability and possibly other aspects.

1.2.1 Development of Playability Heuristics

Even though there is no generally accepted definition of playability, domain-specific heuristic sets have been developed to evaluate video games and their playability. Figure 1 shows an overview of the development history of playability heuristics.

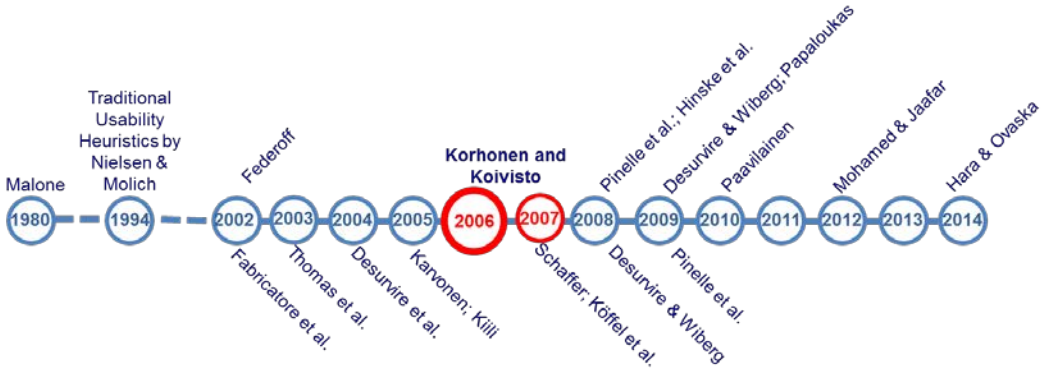


Figure 1. Timeline of playability heuristic development.

The development of playability heuristics was initiated in the 1980s, when Malone (1980) studied video games to find out what makes them enjoyable. He did not call his findings playability heuristics, but they were principles in a design framework. Malone (1980) defined three principles: challenge, fantasy, and curiosity that are needed for designing enjoyable user interfaces. The result highlights the importance of the content in

games. After Malone's work, there have been isolated studies about fun and HCI aspects in games (e.g. Cherny, Clanton, & Ostrom, 1997; Clanton, 1998), but the researchers have not developed their observations further into a form of playability heuristics.

Federoff (2002) conducted a case study in a game company and defined heuristics that can be considered as a first set of domain-specific heuristics for video games. The study did not include the validation of the heuristics, or at least such results have not been published. During the same year Fabricatore et al. (2002) published players in action video games and how their preferences affect playability. They can be used to evaluate action games, even though their findings were not described as heuristics.

Desurvire et al. (2004) published heuristics, which were based on Federoff's heuristics. This was the first study in which the result of the heuristic evaluation was compared to user testing results. The study showed that with the help of heuristics, inspectors are able to identify playability problems in a game prototype. Later on Desurvire and Wiberg published a refined list of heuristics as Game Usability Heuristics (Desurvire & Wiberg, 2009).

It was an interesting coincidence that at the same time as we were working on our playability heuristics for mobile games, there was another study ongoing at the University of Jyväskylä, Finland on the same topic. Karvonen (2005) was exploring factors of playability of mobile games in his master's thesis. The objective of the study was to create a framework of playability from the player-centric point of view and it contained seven categories: 1) Gameplay, 2) User Interface, 3) Audio-Visual presentation, 4) Story, 5) Game Concept, 6) Other factors, and 7) Mobility. We were not aware of this work until it was brought to our attention in September 2006.

The playability heuristics described in this dissertation were originally published in March 2006 (Nokia, 2006). The heuristics were divided into modules and there were four modules called Game Usability, Gameplay, Mobility, and Multi-Player. We expanded the heuristic set in 2008 and defined a module of heuristics for context-aware games (Korhonen, Saarenpää, & Paavilainen, 2008).

After that more domain-specific heuristics for games have been published by different authors. Schaffer (2007) has published a white paper which included illustrations of usability problems related to heuristics. Pinelle, Wong and Stach (2008a) have published heuristics that are based on game reviews. The difference of these heuristics was that there are no heuristics concerning gameplay issues and they can only be used to evaluate game usability issues. Pinelle, Wong, Stach and Gutwin (2009) extended these playability heuristics and published playability heuristics for multi-player games in the next year.

There are also more specific heuristics for certain game genres or player groups. Köffel and Haller (2008) have published heuristics for tabletop games. Educational games and social games have received special attention and there are a couple of heuristics sets or guidelines for these games (Thomas, Schott, & Kambouri, 2003; Kiili, 2005; Hinske, Langheinrich, & Lampe, 2008; Paavilainen, 2010; Mohamed & Jaafar, 2012). Desurvire and Wiberg (2008) have explored inexperienced gamers to give guidelines in how to make games initially more friendly. Recently, Hara and Ovaska (2014) have defined heuristics for the interaction design of motion-controller games. In these games, the user's movements are used to control a game avatar or interact with the user interface. Although these heuristics are targeted to a very narrow scope of issues affecting playability, they indicate the complexity of the game evaluation task.

In addition to heuristics, there are other guidelines available to design more usable and engaging games for players (Falstein, 2003; Snow, 2007). Even though these are not regarded as playability heuristics, they provide useful information for game developers and designers.

Currently, the development of new playability heuristics has settled, and there are several studies which apply the heuristics in different kinds of game evaluations (Röcker & Haar, 2006; Jegers, 2008; Paavilainen, Alha, & Korhonen, 2012). Naturally, there are some special issues which are not covered by the heuristic sets, but the main message in these studies has been that the current playability heuristics cover typical problems in the design and they can be applied regardless of the platform or game genre.

1.2.2 Game Qualities based on Game Critiques

Another favorite method to define good games is to use game critiques to see what aspects game professionals have reviewed in the game. Even though some researchers have used game critiques to define the playability of a game (Karvonen, 2005; Pinelle et al., 2008a; Hara & Ovaska, 2014), they usually take a broader view including marketing and brand issues which are outside the scope of playability evaluation. Bond and Beale (2009) have studied computer game reviews to determine what features should be prioritized to make a game successful (Table 1). They have identified the following 13 categories:

Category	Sample Criteria
Gameplay	Engaging, fair, balanced, progressive, fun, innovative, easy to play, hard to master, objective based, freedom, compelling, dynamic, various possible solutions
Environment	Impressiveness, eye catching, good lighting, lifelike effects, good soundtrack, good sound effects, good music
Storytelling	Mature, progressive, tense, engrossing, embedded in gameplay
User Interaction	Fast feedback, customizable, invisible controls, realistic, functional
Customization	Powerful, easy, personalization, character, modification
Social Interaction	Multi-player co-op, multi-player competition, communication, sharing
Variety	Non linearity, choice, differences, dynamic combat, varied AI, emergent tactics, varied delivery media
Technical Implementation	Well-designed camera, unobtrusive adverts, smooth frame rate, uniformity, freedom to behave as expected
Cohesion	Seamless integration, story related to gameplay, cohesive story, consistent style
Maintenance	Low hardware requirements, easy to maintain, independent of external software
Price (value for money)	Value for money, cost, add on cost, hardware cost
Franchise	Franchise, established genre
Quantity	Lots of good, not much bad

Table 1. Characteristics of successful games by (Bond & Beale, 2009).

Paavilainen (2010) has criticized game critiques as a data source for identifying usability problems. Game critiques have been written for a different purpose and audience, and even though they might report some playability problems, they do not report all of them because game critiques have an impact on the commercial success of a game (Jenkins, Lee, & Archambault, 2010). Zagal, Ladd and Johnson (2009) have studied

game critiques and identified nine themes (Description, Personal Experience, Reader Advice, Design Suggestion, Media Context, Game Context, Technology, Design Hypothesis and Industry) which are commonly discussed in game critiques. Although some of the themes are relevant for the playability point of view, they usually do not discuss aspects on the level that a usability expert would do, but the aspects are left on too high abstraction level. Therefore, using game critiques to determine playability heuristics is questionable.

1.2.3 Heuristic Evaluation Method

Ling and Salvendy (2005) have explored previous work on the heuristic evaluation method and listed research directions that need more work. The research can be divided into four directions that are illustrated in Figure 2.

The domain-specific heuristics research direction is located in the upper part of the triangle because this research direction has been the most active one and several heuristic sets have been developed for different domains. This will help achieve more precise and relevant evaluation results as the original heuristic set cannot cover the characteristics of the domain. (Ling & Salvendy, 2005).

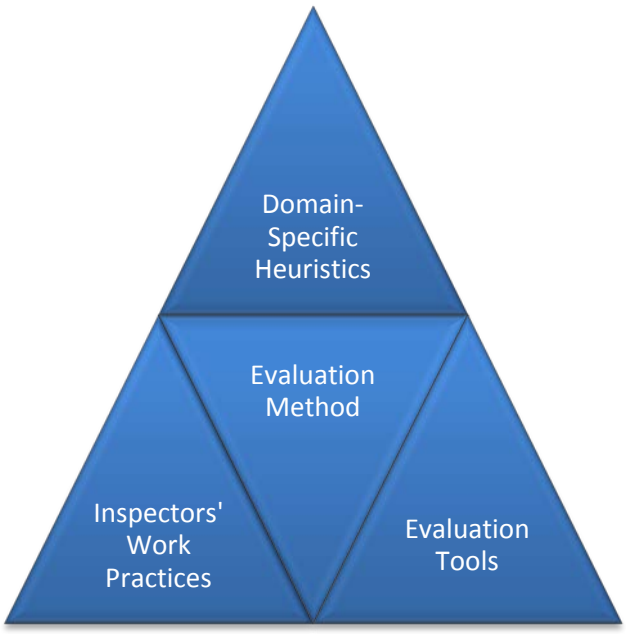


Figure 2. The research directions in the development of the expert review method.

The second active research direction has been the evaluation method development, which is located in the middle of the triangle. This research direction focuses on extensions to the method and how they are applied to various domains. Muller, Matheson, Page and Gallup (1998) have introduced Participatory Heuristic Evaluation (PHE), which includes users of the target system as domain-specific specialists. Sears (1997) combined

heuristic evaluation and another inspection method called cognitive walkthrough resulting in the heuristic walkthrough. Zhang, Basili and Shneiderman (1999) introduced perspective-based methods in which the evaluation session has been divided into several sub-sessions focusing on different aspects. Følstad, Anda and Sjøberg (2010) have studied the performance of work domain experts compared to usability experts.

Another branch in the evaluation method research direction is to focus on different phases of the evaluation procedure. A typical evaluation includes four phases: pre-evaluation training, evaluation task, debriefing phase, and severity rating phase (Nielsen, 1994b). Currently, researchers try to improve the evaluation task, even though improving other phases would be useful as well (Ling & Salvendy, 2005). Some researchers have studied how to improve especially the reporting of usability problems and communicating them to developers (Lavery, Cockton, & Atkinson, 1997; Dumas, Molich, & Jeffries, 2004; Hvannberg, Law, & Lárusdóttir, 2007; Dumas & Loring, 2008).

The Inspector's Work Practices research direction is closely related to evaluation method development, but Ling and Salvendy (2008) state that more research is needed to study the applicability and effectiveness of the extended methods in different domains and how they affect the inspector's cognitive process. They also propose that future research should explore the difference between novice and expert inspectors while they are using traditional or extended expert review methods to improve the method itself (Ling & Salvendy, 2005). This can help in understanding how the different phases of the procedure work and how they should be developed further. Even though the method is intended to be used by experts, who are familiar with the method, the heuristics and the domain they are inspecting, many evaluations are conducted by novice inspectors.

The bottom right corner research direction is evaluation tools development to assist the inspectors to conduct evaluations more efficiently (Ling & Salvendy, 2005). According to my knowledge, this research direction has not been very active and most of the inspectors are still using paper and pencil to record their observations and then transfer them to electronic format when writing an evaluation report of the product. There is probably room for tool development which would help in the evaluation phase and especially in the severity rating phase.

In this dissertation, I focus on three research directions. The first objective is to develop domain-specific heuristics for video game evaluations which are derived from the factors of the games that are important from the playability point of view. There are several studies indicating that domain-specific heuristics are proven to be more effective in evaluations than traditional heuristics because they can capture the characteristics of the system (Baker, Greenberg, & Gutwin, 2002; Berry, 2003; Mankoff et al.,

2003; Somervell & McCrickard, 2004; Zuk, Schlesier, Neumann, Hancock, & Carpendale, 2006). The second objective is to study the evaluation method itself to some extent to find out if game evaluations need some modifications to the original method. The final objective is to study novice and expert inspectors and their work practices in game evaluations to find out how these two inspector groups work in game evaluations.

1.3 Playing Games on Mobile Devices

There are multiple views on defining mobile games. For some people, a mobile game is a synonym for casual games which emphasize simplicity and acceptability in game design (Kultima, 2009) and gaming with mobile devices is seen to happen in passing moments and they are not objects of focused attention (Hall, 2005). Other people think that implementation technologies will define whether a game is a mobile game or not (Järvinen, 2002b). Mobile games may refer to games that extend gaming to the real world environment and mobile devices are used to deliver information between the physical and the virtual worlds (Montola, Stenros, & Waern, 2009). Maybe the most notable difference in mobile gaming compared to traditional gaming is that players do not have to reserve a time and place for playing games, but gaming happens anywhere at any time.

Järvinen (2002b) stated that mobile game design needs to consider the purpose of the device and the context in which the device is used. In this sense, a game that is played on a mobile device is not necessarily a mobile game unless the characteristics of the device and the context are not included in the game concept. The number of such games on the market is not extensive, but smartphones have many technological enablers which open new and interesting possibilities for game design. Game researchers have experimented with these technologies and implemented games which use camera (Hakkarainen & Woodward, 2005; Suomela & Koivisto, 2006), RFID tags (Rashid, Bamford, Coulton, Edwards, & Scheible, 2006), GPS (Global Positioning System) (Björk, Falk, Hansson, & Ljungstrand, 2001; Sotamaa, 2002; Ballagas et al., 2007), Bluetooth radio (Peitz, Saarenpää, & Björk, 2007), Wi-Fi radio (Bell et al., 2006), and different kinds of sensors: accelerometer and gyroscope (Gilbertson, Coulton, Chehimi, & Vajk, 2008) to provide novel game mechanics.

In HCI research, the term *mobile context* is often defined as a single entity in which mobile devices are used (Tamminen, Oulasvirta, Toiskallio, & Kankainen, 2004; Jumisko-Pyykkö & Vainio, 2010). Korhonen, Arrasvuori and Väänänen-Vainio-Mattila (2010) have defined eight context categories: environment, personal, task, social, spatio-temporal, device, service, and access network, which can be used to define the context of mobile product use. According to this, the device is a part of the context and therefore it always influences how the games are played. Bertini et al. (2009) have

concluded that the characteristics of the device such as small screen size, input devices, operating time (battery), network connectivity, computation resources (processor and memory) cause the biggest limitations from the device perspective. Probably the most obvious example of mobile games that use contextual information in game mechanics are location based games (Björk et al., 2001; Sotamaa, 2002; Benford et al., 2004).

As a summary, I would argue that all games that are played on a mobile device are mobile games. Most of the games that we have evaluated during the years could be played on different platforms, but for us they were mobile games because we played them on mobile devices. There are also examples where the same game is available on different platforms. However, playing a game in a certain context with a mobile device will influence the experiences that the game provides for a player.

1.4 Mobile Games on Feature Phones and Smartphones

The era of gaming on mobile phones started in 1997, when Nokia published its first mobile game called *Snake* on the Nokia 6110 mobile phone (Figure 3). The game is probably one of the most played mobile games and it was available on over 350 million mobile phones worldwide (Nokia, 2005b). The game was installed on the device to demonstrate its entertainment capabilities and data transfer speed through Infrared link.

Figure 3. Snake on Nokia 6110 (www.youtube.com/unlokia, 2012)

We started our mobile game evaluations in 2005 and we evaluated mobile games that were played both with feature phones and smartphones. During that time Nokia N-Gage (Figure 4) was a mobile phone, which was specifically designed for mobile games. Its industrial design differentiates this model from other mobile devices and it resembled old handheld game devices. There were also other models such as Nokia 6600, N73, and N81 that were used for playing games. Later on Nokia discontinued the N-Gage product category and transformed it to a mobile gaming service for smartphones. Through an online store it was possible to buy and download games directly on the devices.

The main difference between older smartphones and current smartphones such as iPhone and Samsung Galaxy S is that a touch screen has replaced the keypad as an input device. Instead of pressing keys to control the avatar or navigating in the UI, a direct input method is used and the user

taps directly on the screen. This has enabled bigger screens for devices as the whole A-cover can be reserved for the touch screen. Some playability heuristics presented in Section 5.4 reflect these older smartphones and their industrial design, but as long as these devices are used for gaming, the heuristics can also help inspectors to conduct the evaluation.

Figure 4. Nokia N-Gage Models (Evan Amos, 2014).

The evaluated games represented typical games from various game genres including adventure, action, puzzle, simulation, and strategy (Figure 5). Some games were commercial products developed by 3rd party developers. Nokia provided opportunities for game companies to release their games under the Nokia brand (Mäyrä, 2015). Some games were under development, which provided us with a possibility to give feedback to the developers based on the evaluations and see the impact of the improvements on playability. In order to study the context-awareness of mobile gaming, we developed our own game which included context information to game mechanics.

Figure 5. Mobile games for feature phones and smartphones.

1.5 Objective and Research Questions

Based on the previous work presented in Sections 1.1 and 1.2, it is possible to make some conclusions which define the objective of the study and furthermore, the relevant research questions.

It is somewhat surprising that there are not many studies on playability and its influence on positive user experience. It is especially interesting to note that playability has not been defined even though it has been used extensively in the literature. Some indication of the elements that constitute playability can be seen in the domain-specific heuristic sets that several game designers have designed, but there still are no complete research findings.

Therefore, we can make three observations:

Observation 1: There is no common understanding of what we mean by playability, which will make defining domain-specific heuristics and their coverage difficult.

Observation 2: Games require their own heuristics because the characteristics of games are not sufficiently covered in other heuristic sets.

Observation 3: Utilizing the expert review method in game evaluations is clearly underestimated. Its potential is not fully understood compared to productivity software development projects where it is one of the most commonly used evaluation methods.

When we combine these three observations, the motivation for our study becomes clear. First, there is a demand for a flexible evaluation method that could aid game designers to identify playability problems in the early stages of game development and provide useful information before user testing is a viable option. Second, to achieve effective inspections, we need to have a solid set of domain-specific heuristics that are targeted to evaluate video games. Third, mobile games have become a popular gaming platform which can provide as good a gaming experience as other gaming platforms do.

I will explore the characteristics of playability to find out important aspects that are crucial for a positive game experience and transform these findings into heuristics that can be used to guide inspectors to evaluate mobile games with the expert review method. Nielsen (1993, p. 158) has said that heuristics are common rules that describe common properties of usable interfaces. Similarly, playability heuristics are targeted to describe common properties of playable games.

Even though one of the observations is that usability inspection methods are underused in game evaluation, my intention is not to dismiss play

testing. Instead, as Nielsen has recommended, I would promote the use of both methods to evaluate products, or games in this context, and find as many playability problems as possible. There are several benefits in this approach. Earlier research has shown that both methods tend to find problems which may be overlooked by the other method (Jeffries, Miller, Wharton, & Kathy, 1991; Desurvire, Kondziela, & Atwood, 1992). The inspections can be used to clean up the design from "obvious" errors and, after redesign, user testing can be used to check the redesigned version and find the remaining problems that were not identified by the experts (Nielsen, 1993). This approach will not waste valuable external player resources.

Based on the literature review of the previous studies and observations that can be made, there are three research questions emerging in this topic that I am going to explore in this dissertation.

I. What issues constitute the playability of video games and mobile games particularly?

II. What kinds of playability heuristics would support the analytical inspection of mobile games?

III. What are the benefits and disadvantages of an analytical evaluation method compared to user-based evaluation methods in mobile game evaluations?

In my research work, I will explore these research questions in multiple ways combining the knowledge of the human-computer interaction and the game research domains. The research objective is to develop a playability heuristic set that enables an analytical evaluation of mobile games using the expert review method.

1.6 CONTRIBUTIONS

The research produces three main contributions that address different aspects of my research questions and confirm the findings of previous research:

1. Analytical research produces information that helps define the playability of video games and can be used to develop playability heuristics. It is assumed that this work shows what issues game designers and game researchers consider important when they are talking about playable games. Analytical research also combines important aspects from general usability research in the HCI domain that are common to all kinds of products and should be taken into account in game evaluations, as well.

2. The knowledge and information gained from the analytical research result in a set of playability heuristics that can be used in playability evaluations of video games and mobile games particularly using the expert review method.

3. The series of experimental research produces evidence of how well expert review and playability heuristics specifically work in real game evaluations. The experiments show strengths and weaknesses of the playability heuristics that we have developed when they are compared with other playability heuristic sets and playtesting results.

A tangible contribution of the research is a set of playability heuristics. The playability heuristics cover the user interface aspects which are relevant to all kinds of games. Another highly relevant aspect of playable games is game content, which is also covered by the heuristic set and it is common to all games. For mobile games, we need to cover both the context in which the games are played and the characteristics of the devices. For multi-player games, the social interaction of the players needs to be considered. These aspects give a comprehensive view of the playability of mobile games. The research results help in improving the quality of mobile gaming and entertainment applications on mobile devices in general.

1.7 METHOD REVIEW

This study is based on analytical and empirical research. The analytical part is focused on defining the playability of video games and finding out issues in game design that influence the playability of a game. It is based on a literature review of game design and usability research in Human-Computer Interaction (HCI) areas. The objective was to identify the ontology of games from the playability point of view.

The literature review of game design consists of game design literature that has been published after the year 2000. In addition, I have searched conference proceedings which have mentioned playability as one of their themes. The most notable international conferences are ACE (Advances in Computer Entertainment and Technology), CHI PLAY, DiGRA (Digital Game Research Association), FDG (Foundations of Digital Gaming), and Fun and Games. The literature search included conference proceedings that have been published after the year 2000.

For the HCI research I have searched books, journals, and international conferences. I started the search from the year 1990, when the heuristic evaluation method was introduced. The search contained several digital libraries such as ACM Digital Library (ACM Press), SpringerLink (Springer), Science Direct (Elsevier), and Informa (Francis & Taylor). These

digital libraries include both journals and conference proceedings. In the recent years, some game research papers have also been published in the HCI conferences and journals and they have been included in the literature review.

The primary search terms were *playability*, *usability*, *analytical inspection method*, *heuristic*, *heuristic evaluation*, *expert review*, *evaluator* and *inspector*. Other search terms were also used to find more specific articles.

The empirical part of playability heuristic development consists of a series of mobile game evaluations using both the expert review and playtesting methods. The empirical research includes a) mobile game evaluations using playability heuristics, b) comparison studies in which inspectors used two different heuristics sets to evaluate a game, and c) comparison studies in which playtesting and inspection methods were used to evaluate the same mobile game. The objective was to verify the applicability of heuristics for playability evaluations.

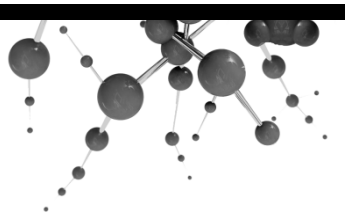
1.8 Structure of the Dissertation

This dissertation consists of a summary and seven original articles published in conferences and journals. The summary will first guide the reader into issues that game designers and game researchers consider important when they are talking about playability in games.

This is followed by discussion of evaluation methods that are used for evaluating games. Special emphasis is on the expert review method and how the procedure of conducting the evaluation differs from normal expert review when we are evaluating video games.

The main contribution of the dissertation is presented in Section 5, which describes the playability heuristics for video games. The playability heuristics have been published in several publications, but due to limitations of publication formats it has not been possible to describe them well enough. This section corrects the shortcoming and describes in detail why each heuristic was included in the domain-specific heuristics for video games and mobile games particularly, and what lies behind the heuristics.

In Section 6 I connect each of the separate publications to the context of the dissertation and explain their role in it. The dissertation ends with conclusions drawn from the studies and discusses the implications that the work hopefully has for practitioners in the game industry.



2 Playability of (Mobile) Video Games

In productivity software, the quality of a product is defined by how well the product meets the requirements of the end-user regarding the tasks a user wants to do with the product (Courage & Baxter, 2005). A usable product supports the user's needs to complete tasks with the product. ISO standard (ISO 9241-11, 1998, p. 6) defines product usability by three attributes.

> Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

The definition says that a product is usable when users can complete their tasks efficiently and they are satisfied with the results. In addition, when this happens in an intended context of use, the product can be considered successful. Nielsen (1993, p. 26) has defined additional dimensions affecting product usability which are related to when a user starts using a product and returns back to the product after some time:

- **Learnability**: The system should be easy to learn so that the user can rapidly start getting some work done with the system.
- **Efficiency**: The system should be efficient to use, so that once the user has learned the system, a high level of productivity is possible.
- **Memorability**: The system should be easy to remember, so that the casual user is able to return to the system after some period of not having used it, without having to learn everything all over again.

- **Errors**: The system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them.
- **Satisfaction**: the product should be pleasant to use, so that users are subjectively satisfied when using it; they like it.

Later on, other researchers have stated that the definition of usability should be expanded to cover emotional aspects of product use. This comes from the notion that functionality and usability of the product alone are not enough to provide user satisfaction (Jordan, 2000, p. 5-6), but the product should provide experiences and pleasures that emerge from the use of the system (Jordan, 2000; Johnson & Wiles, 2003; Overbeeke, Djajadiningrat, Hummels, Wensveen, & Frens, 2003). Carroll (2004) has stated that the definition of usability should be expanded to include aspects that surprise and challenge users. The aspects that attract, capture and hold the users' attention in a certain context make the use of the system fun.

Although it is unlikely that the definition of usability would change, it is interesting to use these notions as a basis when we start defining what playability means in game design. The main objective of video games is to entertain end-users and make the "system" fun to use and provide different kinds of experiences and pleasures. In HCI, emotional aspects have been studied in user experience research (e.g. Hassenzahl, 2003; Bernhaupt, Boldt, Mirlacher, Wilfinger, & Tscheligi, 2007; Desmet & Hekkert, 2007) and they have received more and more attention nowadays.

Game researchers have argued that similar criteria for good usability should not be used when we talk about evaluating the quality of games and other entertainment products. Federoff (2002, p. 8) argues that efficiency and effectiveness are not feasible criteria for game evaluations because they refer to productivity and playing games is an escape from productivity. However, this argument can be challenged because some games require players to work very hard and be productive in developing their avatars to a maximum level and start enjoying the game (McGonigal, 2011, p. 53-54). Further, McGonical (2011, p. 297-298) argues that games can teach us skills (taking a long view, ecosystems thinking, and pilot experimentation) that help us solving real-world problems.

Pagulayan, Keeker, Wixon, Romero and Fuller (2003) have listed the differences between productivity software and games. According to them, productivity software is a tool and the design intention is to make tasks easier, more efficient, less error-prone, and increase the quality of the results. Games, instead, are intended to be pleasurable to play and sufficiently challenging.

2.1 PLAYABILITY ≠ USABILITY

In game design literature, the term playability is rarely used, and in particular, it does not have a similar visibility as usability has in HCI literature. Some game researchers have discussed or mentioned playability (Järvinen et al., 2002; Egenfeldt-Nielsen et al., 2008), but their definitions for playability are mainly abstract, confusing, or otherwise obscure which does not help in providing a clear view of what playability means in games.

For example, Egenfeldt-Nielsen et al. (2008) have stated that during the alpha version of a game, testing is focused on issues concerning ease-of-use and playability. Unfortunately, their definition for playability is rather unexplanatory and contains circular references. It says "a game must be as easy as possible to use (i.e. have high usability) but must of course be fun and appropriately challenging rather than easy to play (i.e. must have high playability, be attractive to play)" (Egenfeldt-Nielsen et al., 2008, p. 19). In this definition, they mainly repeat the basic principle of usability coupled with the fun and challenge requirements of a game. Ease-of-use is both outside and inside the definition of playability. It seems that they have wanted to focus on traditional usability aspects and supplement that with basic requirements for video games. Another confusing aspect in this definition is that high playability of a game means difficulty of playing (as opposed to easy to play). Presumably, they mean that a game should be appropriately challenging, but the choice of words is not very elegant.

Fabricatore et al. (2002) say that playability is determined by a possibility of understanding or controlling the gameplay and any non-functional aspect of the design cannot balance or replace poor playability. Fabricatore's definition is another example of a description that does not describe what playability means, and it is quite limited, focusing only on the gameplay aspect. Snow (2007) has listed basic principles that are relevant from the playability point of view. Control mechanisms that are used to interact with the gameworld should be flexible and customizable and the player should be able to control game flow and narrative elements in the game.

According to Järvinen et al. (2002), playability is a qualitative term which is used for design and evaluation purposes. It refers to guidelines on how to design a game that achieves the desired gameplay or social entertainment. They conclude that playability is a collection of criteria which can be used to evaluate gameplay or interaction. The definition points towards the right direction in saying that playability is a multifaceted issue to be evaluated in games. In addition, Järvinen et al. note that playability could be regarded as a similar research discipline as usability has been in HCI.

Another observation from the work of Järvinen et al. (2002) is that in their definition playability is a four-fold concept which consists of a) functional, b) structural, c) social, and d) audiovisual components. Each of these components is focused on different aspects of gameplay or their influence on gameplay (Table 2).

Playability Component	Description
Functional	Functional playability concentrates on how well the control peripherals and their configuration meet the requirements of the gameplay.
Structural	Structural playability consists of rules, patterns and structures and of how interaction between a player and these aspects happens in a game.
Social	Social playability explores social practices in media use that the product is suitable for in different contexts of use and cultures.
Audiovisual	Audiovisual playability explores the style and appearance of a game and how it relates to the functional and structural components.

Table 2. Components of playability (Järvinen et al., 2002).

Järvinen et al. do not give any detailed definitions for any of these components, but they still give some indication of what kinds of aspects should be evaluated when inspecting the playability of games.

An interesting point by Järvinen et al. (2002) is that depending on the product under evaluation, some of these components might have a bigger role than others which will allow for more flexibility in evaluation. They also recognize the variety of video games and other entertainment products. We have adopted this same principle when defining our playability heuristics. The heuristic set should be flexible and it should support a partial evaluation of focusing only on certain parts of the game or certain aspects of the game design.

Probably one of the most advanced definitions for playability so far can be found in the Usability Glossary (2002) which defines playability as follows:

> 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.

The main message of this definition is that there are multiple aspects that need to be covered by playability. On the other hand, the definition gives a very detailed list of topics that playability covers leaving the definition vulnerable for exceptions or expansions that many researchers can easily think of.

In productivity software, a usable product does not only mean efficiency and effectiveness of interaction and the user's satisfaction. The term has also been used to describe the operating efficiency of the system and the response time of the network (Chandler & Finney, 2005). Similarly, the term playability has been used in some different meanings in media studies or game critiques.

According to Kücklich (2004), game critiques in game magazines relate playability to replayability which means the capability of the game to provide enjoyment for a player over an extended period of time. Kücklich states that, in this sense, playability is an ambiguous term, because replaying the game is influenced by a player's motivation to return to the same game. In his own studies, playability is related to a combination of interaction of a user and the attitude and expectations towards the medium (Kücklich, 2004).

In some cases, game researchers have used the term *game usability* when they are referring to the playability of games (Fabricatore et al., 2002; Isbister & Schaffer, 2008; Pinelle et al., 2008a), and tried to define playability through usability. Federoff considers that game usability is defined by the game interface and playability is a game mechanic (Federoff, 2002). Although game usability sounds like a familiar term and is recognizable by many persons, it may cause confusion or insufficient definitions which merely concentrate on the usability aspects that are specific for games, but it does not specify that there are specific gameplay aspects that are needed in the definition, as well.

Federoff (2002) has found in her study that there are difficulties in understanding the term usability itself among game developers. It is important for usability persons to use language that is understandable to game developers. The lack of terminological understanding has sometimes led to a common misunderstanding of game usability or use of playability heuristics in game evaluations. Pagulayan et al. (2003) have said that the problem of misunderstanding the term is visible in situations where challenging gameplay and a usability problem are mixed up. Games are supposed to be challenging for players, but the challenge should not come from frustrating usability problems that appear in the user interface. The reason for this confusion is that usability guidelines are

only meant for evaluating the user interface, but game evaluations must cover gameplay, as well.

In summary, it is quite interesting that game research literature has not defined playability in detail. Sometimes it feels that playability has been taken into the vocabulary without really knowing what it covers, how it should be understood, or where it should be used. Although there is not an urgent need to have a formal definition for playability as it is done for usability in the ISO standard, it is still useful to explore its dimensions and define the term, because it will then help in understanding, for example, playability heuristics better. This was also a question that we asked when we started our project for developing playability heuristics.

Even though I have complained here on the lack of definition of playability, there are still several indications of issues that can be considered relevant for playability in game design. Game researchers have defined playability heuristics and game design literature is full of excellent books on how games should be designed. These will give hints for defining playability. I will explore some of the available literature further and start formulating a definition of playability which will then be used to define playability heuristics.

2.2 PLAYABILITY INFLUENCES PLAYER EXPERIENCE

The relationship between playability and player experience is similar to the relationship between usability and user experience. Good usability or playability should not be the ultimate goal of the product design, but it is still a necessary step towards positive experience as good usability or playability helps in achieving positive user or player experience of the product or game. Game researchers and game designers have focused on determining player experience and what the components in game design affecting positive player experience are. However, some aspects are more closely related to playability than to player experience.

Federoff (2002) suggested in her master's thesis that satisfaction should have a central role in game evaluations, since the game design objective is to entertain the players. Satisfaction in games could be defined as the games being fun, containing immersive environments and providing compelling experiences for players (Federoff, 2002). Game researchers have not discussed so much about the satisfaction of the players, but they have explored player experience before, during, and after the play session, and say that after the players have played a game, their post-game experience should reflect satisfaction (Poels, de Kort, & Ijsselsteijn, 2007).

From the evaluation point of view, I identified two prerequisites, which are relevant before player experience can be assessed (Figure 6) together

with Elina Koivisto (Nokia, 2006). They are derived from the observations when we conducted game evaluations on mobile devices.

Any difficulty with the stability of hardware or software components will make the evaluation of a game more difficult and possibly hide the potential of the game. The stability of the gaming platform is a foundation on which positive player experiences will be built. Stability means that there are no malfunctions such as crashes and jams which would prevent playing the game, disconnections, or server-side latency. Mulligan and Patrovsky (2003) point out that players will have little patience with these kinds of problems. By a gaming platform we mean the device and the software architecture which is used to play the game. The hardware consists of different components of the device such as processor, memory, display, input devices and other components. The software architecture includes the operating system, drivers, and different engines which are a part of the game mechanics.

The second prerequisite is playability. In order to evaluate the playability of a game prototype, it should include the main game concept and design elements such as a user interface, game mechanics and a view of the gameworld. Finally, when both prerequisites are met, we can start evaluating the player experience before, during, and after the play session.

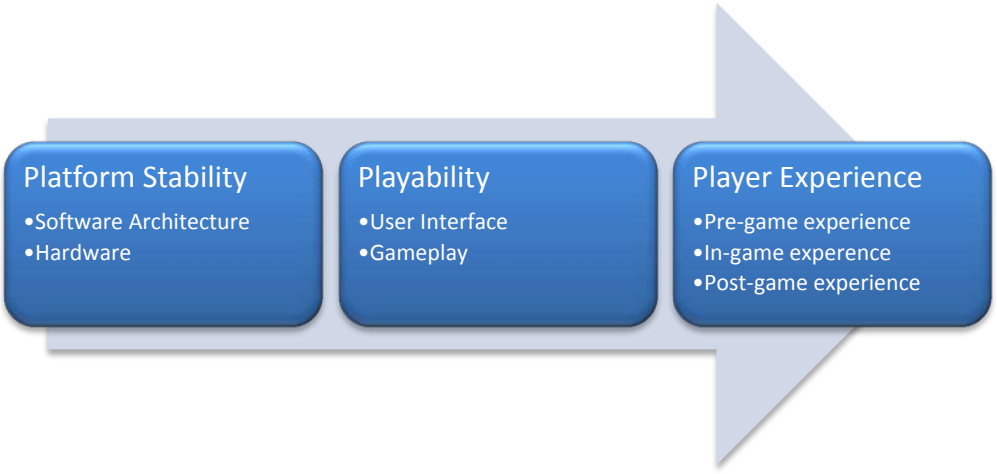


Figure 6. Steps leading to a positive player experience.

There are different kinds of experiences that a player can experience with games. The pre-game experience is generated from the activities that a player takes before the first play session can start. This phase typically includes finding an interesting game and preparing the game for a game session (Kultima & Stenros, 2010). Pre-game experiences have not been studied extensively because they happen before play sessions (Kultima & Stenros, 2010). However, there is some evidence of positive correlation between pre-game activities and player experience. In one study, prior exposure to the background story of the avatar has resulted in a positive influence on the player experience (Park, Lee, Jin, & Kang, 2010). The in-

game experience is influenced by the events that the player confronts during play sessions. Perceived challenge and the story of the game as well as the objectives that the player has in the game will influence in-game experiences. The post-game experience is the result of memories of the game events that the player processes after the play session. Norman (2009) has said that the users' memories of the use of the system are the most important experiences because they will last longer.

Some player experience research gives a more precise view on the aspects that players expect from games. For example, Sweetser and Wyeth (2005) have defined a framework which contains eight components which are critical for enjoyment after gaming. These components are:

1. Concentration
2. Challenge
3. Player Skills
4. Control
5. Clear Goals
6. Feedback
7. Immersion
8. Social Interaction

As we can see, some components such as control and feedback are related to user interface elements, while other components, such as clear goals and challenge, are related to the gameplay. Finally, there are components describing desired states of the player (Immersion) and social interaction as they are playing the game.

Salen and Zimmerman (2004) have used Flow Theory (Csikszentmihalyi, 1991) to describe player experience. They have divided eight components that used to describe optimal experience of a user into two categories which are the effects of the flow state and flow's prerequisites. The interesting part for the playability definition point of view is the prerequisites, because they can be seen as defining factors related to how playability influences player experience positively. These four components are:

- Challenging activity
- Clear goals
- Clear feedback
- The paradox of having control in an uncertain situation.

These four elements need to be designed into games to evoke a positive player experience. Interestingly, all components can be found on Sweetser and Wyeth's list as well. Salen and Zimmerman published their book earlier, but it is unclear whether Sweetser and Wyeth knew about the work.

Salen and Zimmerman have also studied aspects outside Flow Theory and found out that in role-playing games rewards are one of the important aspects that influence player experience. Rewards will influence a player's

motivation and the choices that a player makes in the game (Salen & Zimmerman, 2004). There are four general types of rewards (Hallford & Hallford, 2001, p. 157-160).

1. **Rewards of Glory**: Items that players receive, but they don't have impact on the gameplay.
2. **Rewards of Sustenance**: Items that help players maintain their avatar's status quo and keep all the things they have gained in the game.
3. **Rewards of Access**: The player gets access to new locations or resources that were previously inaccessible.
4. **Rewards of Facility**: Enables a player's avatar to do things they could not do before or enhance abilities that they already possessed.

Although these categories are from role-playing games, they are applicable to other types of games as well, because most game mechanics include some kind of a rewarding system.

Another concept that Salen and Zimmerman promote in good game design is *meaningful play*. It means that goals, challenge and uncertainty of a game provide a context within which choices are integrated and become meaningful (Salen & Zimmerman, 2004).

2.3 MEANINGFUL PLAY

Falstein (2005) analyses meaningful play through a structure and elements of the game and their effects on player experience. The basic game structure divides the challenge and broadness of a game into meaningful choices that a player makes in the game. Schell (2008, p. 179) argues that game design must provide choices that have real impact on what happens next. Falstein (2005, p. 81) cites a well-known game designer Sid Meier who has said that "a great game is a series of interesting and meaningful choices made by a player in pursuit of a clear and compelling goal." If a player believes that one alternative is better than another to reach a goal, then it is a meaningful choice for the player (Falstein, 2005). Schell (2008) says that meaningful choices for a player lead them to ask themselves questions about how they want to play the game.

In game design, it is easy to offer meaningless choices. *Functionally meaningless* choices are often related to the selection of game items. A game might include multiple similar game items, but if they work equally well to accomplish a goal, the selection is meaningless from the player's point of view (Falstein, 2005; Schell, 2008). However, it should be noted that meaningful choices are sometimes subjective and the selection between functionally identical game items can be a meaningful choice, if it

contributes to a player-generated goal, e.g. customizing the game character to look according to a certain style (Falstein, 2005).

The second type of meaningless choice is related to interaction with Non-Player Characters (NPCs). If the player can act differently in the dialogues and either insult, flatter or ignore a game character, but the responses are always identical, then it could be said that the choices are *narratively meaningless* (Falstein, 2005).

An ideal game structure emphasizes the choices a player needs to make and what the consequences of those choices are. The player should start with just a few choices that lead to more (Falstein, 2005). Then the game design starts to narrow those choices back to a single choice the player must accomplish. This core structure is called a convexity because of its outwardly curving shape as shown in Figure 7.

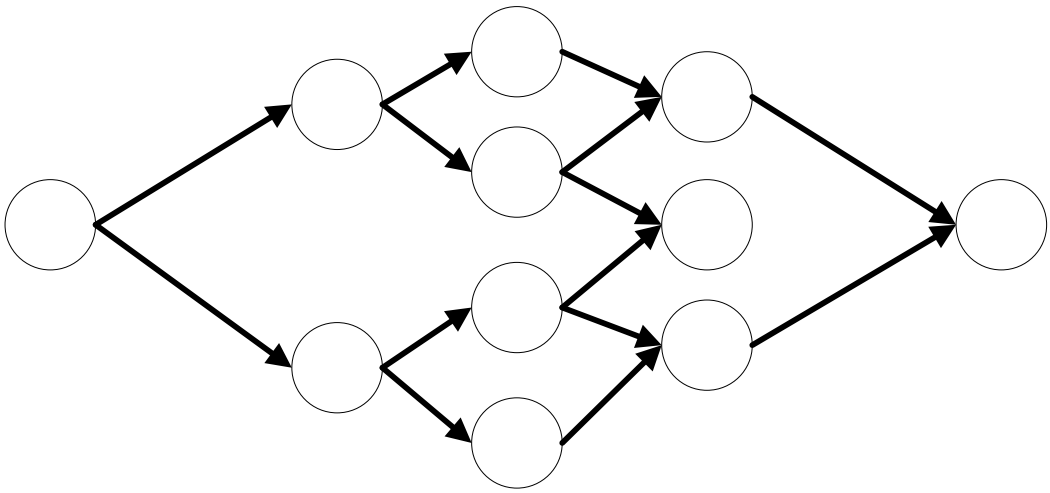


Figure 7. A Single Convexity (Falstein, 2005) Circles represent the possible choices.

The single convexity can be repeated in a game to form a series of convexities (Figure 8) and there are usually also overlapping convexities which will then provide varying paces and challenges for players in the game and provide different tasks to complete at the same time. The layered structure of the convexities ensures that a player always has compelling short-term goals available that provide plenty of meaningful choices to make at all times (Falstein, 2005).

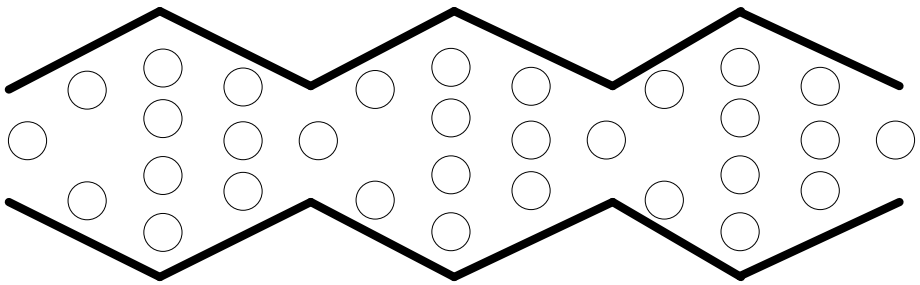


Figure 8. A series of Convexities (Falstein, 2005).

This kind of a structure is effective both from the player's and the game designer's perspectives. For a player, the structure gives plenty of choices where things can happen in any order and provides a feeling of freedom. From the developer's point of view, there are places where you can provide game critical narrative or introduction of new things logically (Falstein, 2005). Even in sandbox games this structure is visible, because there needs to be new game elements which become available as the gameworld develops and unfolds (Breslin, 2009).

The series of convexities will also influence the challenge of the game. It allows the gradual increase of difficulty with each successive level or specific areas with tougher challenges for the players to master (Falstein, 2005). Falstein recommends that although there should be increased difficulty in a game, it is advisable to have the challenge increase sometimes faster and sometimes slower (Falstein, 2005, p. 90). Changing the challenge level will influence the pace and allow the player to learn new skills better and practice them before they are required in the game.

Next I will take a look at how some well-known game designers have pointed out important aspects in game design for creating a playable game.

2.4 THE ELEMENTS OF A GAME

When we look at games, we should be able to identify some basic elements that form a game and are always present in one way or another. The basic elements can be used in the formulation of playability heuristics because they indicate the most important parts of successful games. Several game researchers and game designers have introduced these elements.

One of the well-known models is the MDA (Mechanics, Dynamics, Aesthetics) model by Hunicke, LeBlanc and Zubek (2004). The model only considers the gameplay aspects of the game design and it is used to conceptualize the dynamic behavior of the game system and its impact on player experience (LeBlanc, 2006). Here, aesthetics refer to different experiences rather than an audio-visual perception of the game.

Rollings and Adams break down game design into three areas: core mechanics, storytelling, and interactivity (Rollings & Adams, 2003, p. 8-13). Their model differs from the MDA model slightly because it includes a storytelling part. Interactivity refers to interaction with the core mechanics and the graphics of a game (Rollings & Adams, 2003).

Fullerton et al. (2008) have explored game design from a play-centric perspective. They have identified several common elements that are typical for games. They have divided these into two groups: formal (e.g.

Objective, Procedures, and Rules) and dramatic (e.g. Challenge, Premise, Character, and Story) elements. Although their structure of elements is not a model as such, it can be used to identify possible aspects that should be covered by playability heuristics.

The most feasible model from our perspective comes from Schell (2008, p. 41-42) which was selected for the basis of the playability heuristic definition. Schell's model contains four distinct elements which are interconnected and dependent on each other. In addition, the model consists of elements which take a holistic view of game design. The four basic elements are:

- **Mechanics** are the procedures and rules of the game. Mechanics describe the goals of the game and how players can try to achieve those goals.
- **Story** is the sequence of events that unfolds in the game. Sometimes story refers to a theme of the game, if the game does not have a storyline. A story or a theme is often used to clarify game mechanics to the players and an interesting theme or story will make a game more engaging.
- **Aesthetics** describe how a game looks, sounds, and feels. Aesthetics are an incredibly important aspect of the game design since they have the most direct relationship with a player's experience.
- **The technology** that is chosen for a game enables one to do certain things and prohibits one from doing other things. The technology is essentially the medium in which the aesthetics take place, in which the mechanics will occur, and through which the story will be told.

All these elements are important for game design and they need to be in balance. Further, they all have equally powerful effect on player experience (Schell, 2008, p. 43). For game designers, it is important to understand how these elements interrelate with one another, but it is equally important to consider how they relate to the game experience (Schell, 2008). Bickford (1997, p. 178) has described that the goal is to keep a player "in play as long and as deeply as possible".

2.4.1 MECHANICS

Mechanics are the internal part of a game and a player does not experience them directly. Mechanics are also called the game engine or the physics of the gameworld (Clanton, 2000) because they contain all the data about the game's story, its rules, and its current state (Adams & Rollings, 2007, p. 224). Schell describes six basic mechanics for video games: space, objects (with attributes and states), actions, rules (including goals), skills (of the player), and chance (Schell, 2008, p. 130-169). Mechanics define the

relationship among entities and the events, and the conditions that trigger events and processes (Adams & Rollings, 2007).

Adams and Rollings have presented six tasks that mechanics usually do in a game (Adams & Rollings, 2007, p. 320):

1. Presents active challenges to the player specified by the level design.
2. Receives player actions from the UI and implements their effects.
3. Operates artificial intelligence (AI) of non-player characters and artificial opponents.
4. Keeps track of the current gameplay mode and switches modes whenever the mode change occurs.
5. Transmits triggers to the storytelling engine when in-game events or dramatically significant player actions occur.
6. Operates the internal economy of the game by defining how the game or the player creates, distributes, and consumes the goods on which the game bases its economy.

As a summary, mechanics define how the gameworld operates and how everything in it will behave (Clanton, 1998; Adams & Rollings, 2007).

The critical point from the game design perspective is how the mechanics work together and how they provide a good player experience. There should be a limited number of operative actions for a player in a game. Too many actions, especially those that do not interact with each other well, can lead to a game that is bloated, confusing, and inelegant (Schell, 2008). The players do not know what the mechanics are and can only interpret the functionality of the mechanics from the way the game behaves (Adams & Rollings, 2007).

2.4.2 Game Story and Avatar

Game story is an important part of the player experience and it is used to create a premise (or fiction) of a game (Fullerton et al., 2004) or it can provide a theme that gives a flavor to the game (Falstein, 2005). The story can also direct choices that a player makes in the game (Falstein, 2005).

Game premises can be complex as in the case of character and story-based games. The story can make a player feel that the player is inside the story and is affecting its flow and events (Adams & Rollings, 2007). Alternatively, the story may be little more than an abstract metaphor that allows people to manipulate game items in some understandable way, like arranging shapes in Tetris or stacking colored balls in Magical Drop (Barry, 2005). Depending on the game, the importance of the story changes (Federoff, 2002).

Another objective for the story or the theme is to make activities that are done in the gameworld plausible to a player. A game may contain

unrealistic things, which are needed to make the game more accessible for the players (Schell, 2008). The story can also be used to describe the unusual game mechanics and explain actions and rules that may otherwise be confusing to the player (Barry, 2005; Schell, 2008). Of course, the plausibility of the story is subjective, which creates a challenge for story writers. If the fiction in the game is beyond what a player is willing to accept, then the story fails to enhance the player experience (Barry, 2005).

Game avatars are also an important part of the game and its story. Players and humans in general have an ability to project themselves into an avatar (Schell, 2008, p. 312). This will usually create an emotional link between the game avatar and the player (Falstein, 2005). In a proper emotional state, if something happens to the avatar, the player feels as if it happened to him (Schell, 2008). The projection can happen with any kind of object; human form or an artificial object (Schell, 2008).

2.4.3 AESTHETICS

Game aesthetics have an important role in influencing the player experience and aesthetic considerations make the experience more enjoyable (Schell, 2008). In game studies, there are different interpretations of what game aesthetics mean (Niedenthal, 2009). They can be understood as sensory stimuli that players encounter in a game through visual, aural and haptic sensors. Game aesthetic can also refer to the aspects of digital games that are shared with other art forms and thus, games can be generalized and compared to them (Niedenthal, 2009). Third perspective is to refer to game aesthetics when we are talking about pleasures, emotions and the sociability of games (Niedenthal, 2009).

From the playability evaluation point of view, it is most useful to evaluate the aesthetic aspects as sensory stimuli that game designers have included in a game. Games typically provide visual features as a form of artworks. Beautiful artwork can draw a player into the game and provide artistically pleasurable experiences (Desmet & Hekkert, 2007; Theng, Ho, & Wee, 2008). It can make the gameworld feel solid and real and the player can take the game theme and story more seriously when the artwork supports the game (Schell, 2008). Auditory features have a strong emotional influence on users (Garris, Ahlers, & Driskell, 2002). Auditory features have a similar effect on the game world as graphics. In addition, they are used to provide additional feedback to players in the user interface. As technology has evolved, tactile features start to appear in games (Tan & Jansz, 2008). Tactile features will deepen the visual and aural elements in games.

2.4.4 GAME TECHNOLOGY

Schell (2008, p. 405) talks about *foundational* and *decorational* technologies and the importance of understanding the difference between these two.

Foundational technologies are the ones that make new kinds of experiences possible. Foundational technologies can be new interaction styles, things that improve the usability of a product, a new piece of technology such as a touch screen or a Near Field Communication (NFC) chip on mobile phones, or a new medium type such as CD-ROM and DVD which increased the storage space for game content in earlier days. Decorational technologies just make existing experiences better (Schell, 2008). Schell gives an example of ragdoll physics that make avatars move without predefined animations. This was previously used as decorational technology, but it can also be seen as a foundational technology in many games.

2.4.5 The Interaction between Game Elements

The main point of playability evaluation is to evaluate the interaction between a player and the game elements. Mäyrä (2008, p. 17) presents that every game has a core (gameplay) and a shell (symbolic presentations) that operate on the gaming platform. The whole structure of the game elements (Figure 9) needs to be considered when evaluating and analyzing games. Each of them is a possible source of playability problems, which are either seen directly or through another game element.

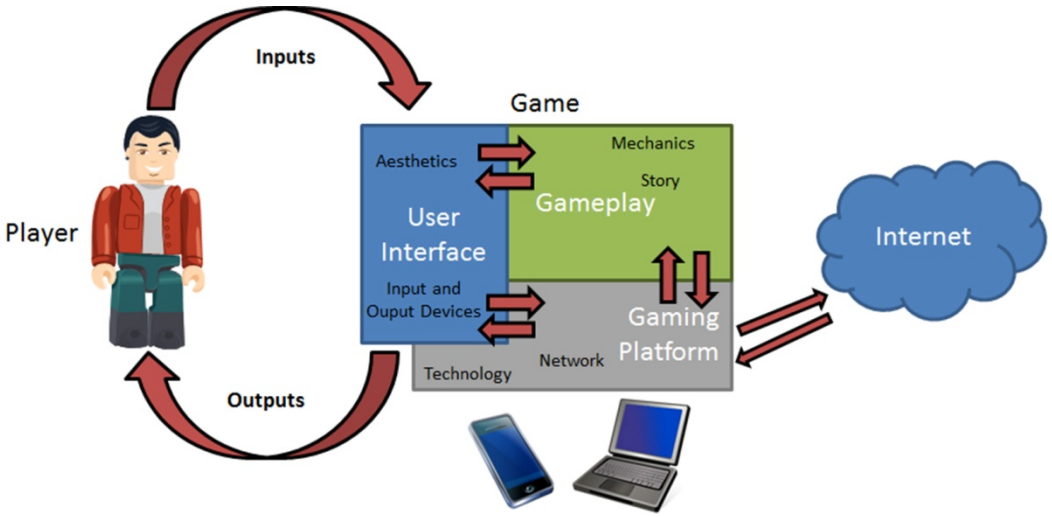


Figure 9. The interaction between game elements.

The player interacts with the game mechanics through a user interface, which plays a crucial role for providing immersive interactivity for a player (Schell, 2008). It consist of input devices to control the game avatar, and output devices that provide game information for a player (Clanton, 2000). The input and output devices can be either software or hardware components. Sanchez-Crespo Dalmau (1999) states that interfaces are considered best if the player does not notice them or they are considered otherwise invisible. In an immersive environment, a user should forget that they are participating through a medium (Federoff, 2002). Although the interface is not the main aspect of user satisfaction in games, a poorly

constructed interface can keep a player from enjoying the gameplay (Shelley, 2001).

Gameplay interacts with both the user interface and the gaming platform. The most notable parts of the gameplay are the mechanics and the story. Gameplay includes problems and challenges that a player faces and wins in a game (Federoff, 2002). Scoring and rewards are considered to be a part of gameplay (Clanton, 2000). Gameplay also defines the pace and the cognitive effort that the player needs to make. Gameplay receives input commands from the player through the user interface and the gaming platform and processes them in the game engines. The results of the actions are presented to the player using output devices.

The gaming platform enables interaction between the player and the gameplay. The gaming platform consist of the available processing power, memory, a graphical accelerator, network connection or other features of the platform. Sometimes limited resources can influence the gaming experience. If the game, for example, consumes too many resources to present the current graphical settings of the game or the gaming platform is busy in processing some other tasks, it will cause delays in the gameplay and create frustration for the players. Another important task is to maintain the connection between the gaming device and external game servers or other devices. Many games are not standalone products anymore, but include social interaction which enables interaction with other players. The stability of the network connections and speed are essential factors forming the player experience in multi-player games.

2.5 Components Forming the Playability of a Game

In the previous section, we discussed basic game elements (User Interface, Gameplay, and Gaming Platform) and how a player interacts with these elements. Playability issues usually emerge if either game elements are confusing or unclear for the player or the interaction with the game elements turns out to be difficult. All game elements should be covered in a playability inspection. It is evident that playability should be understood as a broader term than usability because it needs to cover multiple aspects of a game. Moreover, multiple elements implicate that playability should be defined by using a multifaceted description.

The playability of a game is related to challenge, fun, intuitiveness, and unobtrusiveness. Social interaction in multi-player games and possibly some other aspects will also affect playability. Based on the three components presented above, we define playability in video games as follows:

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.

This definition will help us start defining playability heuristics because now we have topics that should be covered in the heuristics. In the following sections, we will describe each of the components and look in more detail at specific aspects which affect playability. Playability heuristics and their descriptions are presented in Section 5.

2.5.1 THE GAMING PLATFORM

The gaming platform is the component through which a player interacts with the technology elements of a game. The game designer usually does not have much control over the gaming platform, but the design must be adjusted to the capabilities of the platform.

There are three typical platforms which are used for playing video games:

- On **game consoles** games are played with dedicated controllers which have a limited number of buttons. The games are played in front of the TV screen (preferably HD televisions).
- **PC** games most commonly use standard input devices such as a mouse and a keyboard. The output device is a monitor which has a good resolution. It is also possible to use dedicated controllers such as a driving wheel for some games, but most games are played with a mouse and a keyboard.
- On **mobile phones**, a player uses input devices that are integrated into the device itself. It can consist of an ITU-12 keypad, a touch screen, or a full QWERTY keypad with remarkably smaller keys than on a normal keyboard. In early days, there was a joystick mounted on top of the mobile phone keypad to make playing on mobile phones more convenient (Figure 10). Today, touch enabled devices have become mainstream and this allows direct manipulation of objects on the screen.

Figure 10. Joystick mounted on a mobile phone (http://mobil.idnez.cz, 2001).

Although input and output devices are included in the user interface component, they have a relationship to the gaming platform as well because the gaming platform defines its typical input and output devices. The game designer needs to take the characteristics of the input and output devices into account and it is advisable to design for the default control devices. Specialized control devices should be used only when they significantly enhance the player experience or the game intentionally exploits new technologies (Adams & Rollings, 2007).

2.5.2 THE USER INTERFACE

A game can have innovative gameplay, astonishing artwork and story, but a smooth and intuitive user interface completes the player's perception of the game (Adams & Rollings, 2007). The user interface is also called the presentation layer as it makes the game visible and audible to a player (Adams & Rollings, 2007).

The user interface implements two important aspects of a game: interaction model and perspective (Adams & Rollings, 2007). The interaction model determines how the player interacts with the gameworld. The user interface delivers the player's commands from the input devices and interprets them as actions in the gameworld passing those actions to the mechanics (Adams & Rollings, 2007). Perspective determines how the player sees the gameworld and how the camera behaves (Adams & Rollings, 2007). There are a few main perspective modes that are used in most games (Järvinen, 2002a).

The user interface also presents data from the game mechanics to the player in visible and audible forms (Adams & Rollings, 2007). The challenge of user interface design is to select which data is presented in the user interface and how game state changes are communicated to the

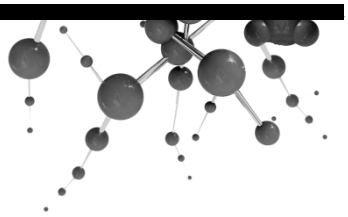
player. Some game states can be public, limited visibility, private or random. Games that force players to be aware of too many states can confuse and overwhelm the players (Schell, 2008).

Game user interfaces are unique compared to other software applications. The main difference is that most software applications are tools and usability criteria of such tools is to allow the user maximum control, information visibility and flexibility of doing tasks. In addition, the user should be able to enter and create data and clearly see the results (Benyon, Turner, & Turner, 2005). In video games, the user interface should also provide control, information visibility and flexibility, but not to the same extent as in productivity software. There are lots of things happening in the mechanics that are hidden from the player. The user interface mediates between the internal parts of a game and the player, creating a player experience for the player which communicates the gameplay and storytelling in an exciting and engaging manner.

Another difference between the user interface of video games and other software applications is that for video games user interface design rarely follows any established conventions and every game usually has its own user interface style. Game designers are urging innovation in user interface design because innovation is greatly valued in almost all aspects of game design: theme, gameworld, storytelling, art, sound and gameplay (Adams & Rollings, 2007).

2.5.3 GAMEPLAY

Gameplay is the core of what a game truly is (Schell, 2008). Gameplay is also the most determinant feature influencing the playability of a game (Karvonen, 2005, p. 153). Gameplay is about the challenges and actions the players take to achieve goals and get rewards that exist in a game. Narrative is used to explain the environment and the things that happen in the game. All gameplay aspects are hidden from the player at first and revealed piece by piece as a response to the player's actions. This makes games unique compared to other software applications.



3 Evaluation Methods

Product designs are typically evaluated by using user-based testing and inspection methods. In the former method, participants from the target user group are brought to a usability laboratory and their performance is observed during a task completion. In inspection methods, the evaluation is conducted by experts who go through the product systematically and report identified problems.

Cockton and Woolrych (2002) questioned whether products should be evaluated using either a user testing or inspection methods, but the fact is that neither of the methods is better than the other and both types of methods are useful for their own purposes. Preece et al. (2007, p. 427) also cite Nielsen as stating that usability inspection methods should be used together with user-based testing to get more comprehensive evaluations.

In game evaluations, user-based testing is dominating and game design literature describes playtesting as a preferable method for evaluating games (Rouse, 2001; Fullerton et al., 2004; Schell, 2008). However, HCI researchers have concluded that usability inspection methods can provide valuable feedback to designers (Molich, Kaasgaard, & Karyukin, 2004; Dumas & Loring, 2008), and they can be more agile during a game development project.

3.1 Usability Inspection Methods

The interest for developing evaluation methods that do not involve users from the target user group was initiated in the 1990s as there was a need to cut the costs of usability testing. One of the first workshops about the usability inspection methods was organized at the ACM Human-Computer Interaction conference (CHI) in 1992 by Robert Mack and Jakob

Nielsen (1993). The first two inspection methods were heuristic evaluation and cognitive walkthrough (Nielsen, 1994b). Since then, the set of inspection methods has been expanded and eight methods [1] can be considered as inspection methods: *heuristic evaluation*, *guidance review*, *pluralistic walkthrough*, *consistency inspection*, *standard inspection*, *cognitive walkthrough*, *formal usability inspection*, and *feature inspection* (Nielsen, 1994b). In addition to these methods, there are also variants for almost each of them having aspects that were not originally included in the method.

The inspection methods are well-known and heuristic evaluation is the most popular method to evaluate the usability of products. For this reason, heuristic evaluation was selected to explore its suitability for evaluating games.

3.1.1 The Benefits of Usability Inspection Methods

The original objective of developing the heuristic evaluation method was an attempt to reduce the costs of evaluation. For this reason, the method has also been called a discount usability evaluation method in the literature. User testing is usually considered expensive and time-consuming. Many previous studies have concentrated on proving how much usability evaluation can reduce the costs in different contexts (Bias & Mayhew, 2005).

Another benefit of the method is that it is less time-consuming and requires less effort than user testing. An evaluation can be conducted in a few hours and the results are often reported to the developers within the same day. The method can also be used more frequently to evaluate revised versions of the design. This allows for agile product development, which is typical for game development projects (Clinton, 2010).

The third benefit of the method is that the evaluation can be done with many different kinds of prototypes. A low fidelity prototype can include only partially implemented features or the evaluation can focus on a specific task of the user. In the earliest phase, heuristic evaluation can be used for evaluating the design documentation which only contains sketches of the planned user interface. In game evaluations, gameplay could also be evaluated by using videos of the gameplay or watching someone else play it. The choice of the inspectors is the only limit in this sense.

3.1.2 The Challenges of Usability Inspection Methods

Using usability inspection methods has also drawn some criticism and sometimes study results have been questioned. Soon after the heuristic

[1] A short description of the inspection methods can be found on http://www.useit.com/papers/heuristic/inspection_summary.html

evaluation method was introduced, several researchers conducted comparison tests to evaluate the effectiveness of different methods (e.g. Jeffries et al., 1991; Desurvire et al., 1992; Karat, Campbell, & Fiegel, 1992). However, later studies have found several problems in the experiment design and the conclusions drawn from the results (Gray & Salzman, 1998). Some researchers have even suggested that "discount methods" whether they are inspection methods or user testing should be removed from the HCI method toolbox completely because they try to reduce the demand on resources, time and skill that are required to effectively conduct an evaluation (Cockton & Woolrych, 2002). Understanding usability problems requires attention on three things a) the context in which they arise, b) the actual immediate and eventual difficulties of the user, and c) the assumed causes of these difficulties (Lavery et al., 1997).

3.2 EXPERT REVIEW

Nielsen and Molich (1990) developed the heuristic evaluation method. A small inspector group examines the product and its interface and determines compliance with the help of usability heuristics.

Heuristics are "a commonsense set of rules intended to increase the probability of solving some problem".[2] Heuristics are used as an aid to serve learning and discovery, or problem-solving by experimental methods[2]. In usability evaluations, heuristics focus on issues that cause problems during interaction. The first version of usability heuristics was published together with the method, and the revised version of the heuristics was published in 1994 (Nielsen, 1994a).

Nielsen has described that the evaluation determines good and bad aspects in the interface based on the heuristics (Nielsen, 1993, p. 155). Simply inspecting the interface originates from the practice that the method can be used to evaluate non-working prototypes such as paper mock-ups which do not enable real interaction with the interface or the system (Nielsen, 1990). Inspection results include identified problems with suggestions on how to fix the problems and they are reported in such detail that the development team can come up with an alternative solution.

On many occasions, the method is called expert review (Jacobsen, Hertzum, & John, 1998; Zazelenchuk, 2006). Rolf Molich also recommended[3] calling the method expert review rather than heuristic evaluation. Expert review can be seen as a more appropriate name for the method because, in practice, the evaluation is rarely conducted in a manner that Nielsen and Molich originally intended. The inspectors, especially when they are experts, have learned the heuristics and are

[2] http://www.webster-dictionary.org/definition/Heuristic
[3] Personal communication with Rolf Molich, 2006

familiar with usability principles in different product categories, which will give them the expert status. In addition, the inspectors will not only rely on heuristics when they inspect the product, but their own expertise of conducting evaluations and a previous knowledge of similar products will also affect the quality of the evaluation.

In game evaluations, the method should be called expert review rather than heuristic evaluation because in a game evaluation the inspectors' experience of different game genres, their gaming experience and gaming skills have an essential role in the evaluation. The inspectors must play a game in order to evaluate the gameplay. Of course, some parts of the user interface, such as the shell menu, can be evaluated by inspecting the design, but the goodness of the user interface will become evident once it is tested during actual gameplay.

3.3 A Procedure for Conducting an Expert Review for Games

The description of the method and the procedure for conducting an expert review have been described by Nielsen in two books. *Usability Engineering* (Nielsen, 1993) is a complete book of usability engineering practices starting from a general definition of usability, usability engineering lifecycle, and then proceeding to describe how usability can be evaluated during product development. In the chapter about usability heuristics, Nielsen gives a thorough explanation of 10 heuristics that are used in many evaluations. *Usability Inspection Methods* (Nielsen, 1994b) focuses more on different inspection methods that can be used to evaluate products when the evaluation is conducted by experts or inspectors. In this book, heuristic evaluation is one method among others. The chapter focuses more on the procedure of conducting the evaluation and introduces new aspects to the method.

The procedure of conducting an evaluation is described in a similar fashion (sometimes literally) in these two books. Other researchers have proposed some modifications to the method (e.g. Kurosu, Matsuura, & Sugizaki, 1997; Sears, 1997; Chattratichart & Brodie, 2002; Po, Howard, Vetere, & Skov, 2004). The main idea has been to divide the evaluation session or the heuristics into smaller units to make the evaluation more efficient.

The following is my summary of the procedure which is based on the procedure presented in those books. In addition, I describe differences in the procedure of game evaluations. These observations stem from the evaluations of mobile games that we conducted when developing the playability heuristics. Table 3 summarises the differences of the evaluation procedure between productivity software and game evauations.

	Productivity software	Games
Evaluation rounds	2 rounds	3 rounds
Task scenarios	Sometimes for specific domains	Rarely
Problem identification	Linear progression	Multipath progression
Benefits of secretary	Provides assistance in specific domains	Records identified problems
Duration of evaluation session	1-2 hours	2-4 hours
Consolidated problem list	Single perspective – User Interface	Multiperspective – User Interface vs Gameplay
Inspectors	Double Experts Recommended	Double Experts Mandatory

Table 3. Summary of differences in the evaluation procedure.

3.3.1 A GAME EVALUATION LASTS FOR THREE ROUNDS

Nielsen recommends that inspectors should go through the interface twice. The first round is intended for getting an overall feeling of the interface and of the interaction flows when users perform tasks. The second round is used to focus on some specific parts of the interface. Naturally, the inspectors can decide on their own preferred procedure of exploring the user interface, but going through the interface twice has its advantages.

In game evaluations, the procedure of conducting the evaluation should be divided into three rounds. The first round is dedicated to exploring the interface elements which are outside the actual game. This usually includes the shell menu, the settings and other supportive interfaces. This task is similar to productivity software evaluations.

In the second round, the inspector should evaluate the gameplay by playing the game. Capturing things that the inspectors encounter for the first time is a critical moment because players will learn and adapt quickly to the gameplay. If there are playability problems, they should be recorded immediately, otherwise they might be missed. Evaluating the gameplay is the most challenging part of the evaluation because the inspectors have to develop the required skills for playing the game and be able to proceed in the game within the time limits of the evaluation. The gameplay evaluation is about exploring goals, challenge, the gameworld and other aspects which are parts of the game. The inspectors need to

know how much control is given to a player and how difficult the challenges are in the game.

In the final round, the inspector should examine the game interface through which a player interacts with the gameworld and the gaming platform. Getting familiar with the game mechanics, objectives and goals before evaluating the game interface and the gaming platform is an important step in order to be able to judge how well the game interface supports playing the game and whether it provides accurate and sufficient information to the player. Most game genres have evolved from early version to state of the art games including a practical set of feedback elements and control mechanisms suited to the genre. Although innovation is highly appreciated in the game industry, it will also bring risks from the playability point of view. A completely new interface style will increase the learning time of the game. Forcing a player to learn an unfamiliar user interface will most certainly frustrate the player and reduce enjoyment of the game (Adams & Rollings, 2007). A novel user interface style can differentiate a game from other similar game concepts and make something familiar look completely novel and different. However, it may lead to an increased risk of introducing design solutions that may contain playability problems and overlook previously successful designs. Therefore, the inspectors need to be aware of game genres and previously published game titles of the same genre.

The main objective of the gaming platform evaluation is to see whether the platform is feasible for the game and able to run the game smoothly. The player should be able to control the avatar according to the requirements of the gameplay. Sometimes the game mechanics may require too much processing power that some gaming platforms (e.g. certain mobile devices) are not capable of running the game fast enough. Similarly, the gameplay can be so fast that a touch screen based interaction model is not reliable enough for fast and accurate control of the avatar.

3.3.2 Task Scenarios Are Seldom Used in Game Evaluations

In productivity software evaluations, inspectors are free to explore the interface as they like, but in some cases it might be necessary to supply the inspectors with typical task scenarios or a list of steps of how a user would perform tasks within the system. Such supplemental material might be needed in highly specific domains which the inspectors are not familiar with. The usefulness of such material has been studied by Carrol and Rosson (1992) and Clarke (1991). Nielsen notes that if the task scenarios or any other instructional material is used, it should be based on a task analysis of the actual users and the tasks they perform with the system in order to be representative examples of how the system is used.

In game evaluations, task scenarios are rarely needed. Especially at the beginning of a game, there is usually a separate introduction session or a

tutorial that introduces the player to the game and gives instructions on how to start playing the game and what to do in certain situations (Mulligan & Patrovsky, 2003, p. 138). In some cases, the inspectors may be assisted by providing primary and secondary goals that players try to achieve in the game. This will help the inspectors to examine the most important parts of the game.

3.3.3 Identifying Playability Problems

Identifying usability problems in a product can be categorized into three groups. The easiest group of problems is those that are located in a single instance of the user interface and are spotted once they are encountered. Typically these problems are related to navigation or layout. The second group of problems is those that can be located in two or more locations in the product. This requires that the inspectors cross-check the user interface to find problems. These problems are typically related to the consistency of the user interface. The most difficult group of problems is those that are related to missing features in the design. Nielsen has commented that the missing features are usually identified once the inspectors get stuck and would need something which is currently unavailable in the interface (Nielsen, 1994b, p. 56-57). In productivity software evaluations, there is no big difference between how frequently the inspectors have identified different kinds of usability problems (Nielsen, 1992a).

In game evaluation, consistency problems are challenging to find because the gameworld creates one big complex space for interaction and the players should be able to do similar kinds of things throughout the game. Encountering problems due to missing features or incomplete quests in earlier stages of a game is also a challenge for the evaluation. Many games do not have a linear path to progress and the players can end up in situations where the game is assuming that the player has certain knowledge, skill, or a game item, which is needed for progressing further in the game, but the player does not have them because he has selected another path in an earlier game stage.

3.3.4 Conducting the Game Evaluation

Playing video games and particularly mobile games can happen in many different places. The gaming platform, whether it is a PC, a game console or a mobile device, will either enable or restrict the place where the game is played. Playing games with a desktop PC usually happens by sitting on a chair in front of a desk. Game consoles, mobile devices and laptops are not so restricted in a similar way, but the players can be sitting on the floor or a comfortable armchair while playing the game. The inspectors of game evaluations seem to take the same freedom and evaluations happen in the most convenient place as long as there is room for the heuristics printed on a paper sheet and a notebook.

Using the heuristics does not differ significantly in productivity software and game evaluations. During the evaluation, the heuristics are next to the inspector so that they can be easily scanned through while playing the game. This will help the inspectors to keep them in mind and provide a quick help when needed. It is preferable if the inspectors knew all the heuristics by heart because then their use is more effective. The inspectors are instructed to write down the description of the observed problem by using a problem report (Figure 11), and to report one problem per report. Sometimes the inspectors will get a generic problem criterion such as "[i]s there a playability issue that disturbs your gaming experience in the game?" (Korhonen, 2011). This reminder will especially help novice inspectors to report any issue that they find disturbing, because, in game evaluations, things happen unexpectedly and the inspector needs to be alert all the time. Especially playability problems related to difficulty and the pace of the game, the balance of the game items, and control of the avatar are things that need to be observed constantly.

Evaluator ID: _____

Finding No: _____

Identified Playability Problem:

_____
_____
_____
_____
_____
_____
_____
_____

Write down a short description of the problem or a specific issue. The found problems or issues are features that affect the playability of the game. Every game also contains good things that should not be removed or changed.

Violated Heuristic: _____ (e.g. GU2 or 2)

Severity of the Problem: _____ (Low, Medium, Critical)

Critical =	Occurs every time and it is unavoidable. It affects greatly to gaming experience. Behavior of the game differs from how the player expects it to be
Medium =	Annoying, but the player can live with it, the player can recover from the problem quite easily or avoid it
Low =	Noticeable, but it does not prevent the player to accomplish the goals. Sometimes these are just cosmetic problems

------✂----------

Figure 11. Playability problem report.

The inspector describes the problem using his or her own words and using common terms and sentences. It is not important to describe the problem in full detail at first, but to write down details which are enough to locate the problem afterwards. Usually the inspectors need to recheck the observed problem and play the game again verifying if there is a problem or not. Identifying and rechecking playability problems can happen one after another. Once the problem is verified, the inspector will write a complete report and assign a violated heuristic to it. The heuristic has an

important role here because in the rechecking phase the heuristics can be used to evaluate the problem and find the roots of the problem. Finally, the severity rating of the problem is defined for prioritizing the playability problems for product development. Writing a suggestion for correction is not mandatory at this phase, because it will be done after the prioritization.

3.3.5 A Secretary would Help in Game Evaluations

Reporting the identified usability problems can be completed in two ways. Usually each inspector is responsible for reporting their own findings as written reports which are then aggregated. The written reports have the advantage of presenting a formal record of the evaluation (Nielsen, 1993, p. 157).

The second practice is to have a secretary present during the evaluation session. With this approach, the inspectors only need to verbalize their observations which are then recorded by the secretary. Using a secretary reduces the workload of the inspectors and allows them to use more of their time to evaluate the interface. In some evaluations, a secretary is used to speed up the process and have the results available sooner after the last evaluation session (Nielsen, 1994b, p. 41). In this case, the secretary records the identified problems from multiple inspectors simultaneously and the secretary needs to prepare his or her own notes and not go through all reports from every inspector. The disadvantage of this approach is that every identified problem from the inspectors may not be recorded.

In game evaluations, a secretary can benefit the inspectors even more due to the special nature of the evaluation. Playing the game can be very intense and pausing the game and writing notes in the middle of the game will interrupt the flow and break immersion. It is more convenient to just verbalize the problems and a secretary will write them down. After the play session has ended, the inspector can go through the notes with the secretary and supplement the problem descriptions with more details. Of course, in slow-paced or turn-based games that have natural breaks, the inspectors can write problem reports themselves while playing the game.

Sometimes in productivity software evaluations, the secretary can assist the inspectors in operating the interface, especially in the case where the system is highly domain specific or there is a need to fix problems with an unstable prototype. In these cases, the secretary is someone from the development team.

The secretary can also answer questions the inspector may have and provide hints in difficult situations. However, it should be noted that assisting the inspectors or giving hints should happen only after the inspector has clearly stated a possible usability problem in the interface and it has been recorded. This order is necessary to ensure that all issues

causing problems are recorded and any valuable information is not lost. Providing assistance to the inspectors can be justified with the fact that it is not feasible to waste the inspectors' time by letting them struggle with the interface and not helping them out.

In game evaluations, it should be carefully thought when the inspectors require assistance because the main objective is to evaluate the gameplay which purposely contains difficult things. The inspectors will evaluate the understandability of goals, challenges and the current difficulty of the gameplay that players face in the game. Giving assistance without spoiling the gameplay inspection is quite difficult.

Another possible method of assisting the inspectors is to give them cheat codes for the game (e.g. Laitinen, 2006). Cheat codes are typically used by quality assurance personnel when they want to test certain game features and find bugs in the implementation. The cheat codes will typically enhance the avatar and enable faster progress in the game. However, it is not advisable to use cheat codes during a playability evaluation because the nature of the evaluation will change and the results will not be reliable anymore. Especially gameplay problems may be overlooked if the avatar is somehow better than what it normally would be.

3.3.6 The Duration of the Evaluation Session

In productivity software evaluations, the evaluation session will typically last for one or two hours at maximum. Naturally, it is possible to have longer sessions, especially with very complicated interfaces, but in these cases it would be better to split up the evaluation into several sessions, each concentrated on individual aspects (Nielsen, 1993, p. 158). Productivity software evaluations can be complete evaluations, meaning that the inspectors are able to go through every aspect of the interface during the evaluation session.

Game evaluations are hardly ever complete and the duration of the evaluation session is usually more than two hours. This comes from the fact that playing a game is more demanding than simply going through each interface element. The inspectors need to solve the challenges the game provides in order to progress in the game which takes time. Usual playing time for many games is tens of hours. However, the complexity of a game is not a big problem since the gameplay and the user interface remain the same even though players face new challenges and game content. The main thing the inspectors need to be aware of is the possibility that the game becomes too repetitive and does not provide enough challenge or varying or interesting game content for the players which is regarded as a playability problem. The evaluation can be considered finished when new issues cannot be found easily or the time required to find playability issues is increasing dramatically. In our studies, a single evaluation session has lasted for two to four hours.

3.3.7 THE CONSOLIDATED LIST OF PLAYABILITY PROBLEMS

The primary outcome of expert review is a consolidated list of usability problems that the inspectors have identified during the evaluation session. The list is generated based on discussions between the inspectors in a debriefing session and it reflects the common understanding of the problems the user interface has.

Each usability problem is justified from different angles. The problem description should identify the location of the issue, why the issue was determined to be a problem and how it could be corrected. In addition, the usability problems should always be annotated with a reference to a heuristic or any other usability principles that the design violates. The objective is that each problem is broken down so that one heuristic per problem is used as a reference.

In game evaluations, a consolidated list of playability problems is generated in a similar way and it is based on the inspectors' reports. There are still several challenges to generating a consolidated list of playability problems.

The main challenge of the consolidated list is to determine whether two playability problems are actually describing the same problem or are two individual problems which appear on the same occasion. Especially gameplay problems should be interpreted with caution. Sometimes it may seem that a playability problem has been reported as a user interface problem, but it refers to a heuristic from the gameplay module. In this case, the analysis of the problem should be continued to determine the origin of the problem and assign a violated heuristic accordingly. This is one of the main differences compared to productivity software evaluations because in these evaluations only the user interface aspects are evaluated.

Another challenge of the consolidated list is to determine the location of the playability problems. Playability problems which are related to gameplay can appear in many different locations depending on the player progression and previous accomplishments in the game. For example, the perceived difficulty of a game will depend on the level of the avatar and the number of tasks a player has completed before the current task.

The third challenge is in prioritizing the identified problems. Normally the consolidated list contains problems related to both the user interface and the gameplay. Gameplay problems are typically more difficult to correct and corrections will usually cause a chain reaction to different parts of the design, and thus, are more laborious to correct. As expert review typically produces a large number of usability problems and due to time restrictions and other reasons it is not possible to fix all of them, designers need to decide which problems to fix and which problems to leave intact.

Suggestions for correcting the problems can also be generated until this point because after prioritization it is known what will be corrected.

In game evaluations, there are also different types of playability problems which affect different parts of the game: the user interface, gameplay, mobility or social interaction. According to Nielsen (1993) the severity ratings can help in prioritizing resources to those problems which have the most impact on usability (and playability). However, the severity rating should be based on ratings from multiple inspectors since ratings from a single evaluator are not reliable enough to base any major investments of development time and effort. In practice, there should be ratings from three to four evaluators and by calculating the mean severity rating would give enough accuracy for the decision making.

The severity of the problem is a combination of four aspects:

a) The frequency of the problem occurrence to the players.
b) The impact of the problem on the players' performance.
c) The persistence of the problem, is it avoidable or repeatedly encountered?
d) The popularity of a problem in relation to market impact, even though it is quite easy to overcome.

Instead of listing all of these aspects separately, it is common to present a single severity rating as an overall assessment of the problem. Nielsen (1994b, p. 49) has used a five-point scale for the severity ratings of usability problems (Table 4):

Value	Description
0	I don't agree that this is a usability problem at all
1	Cosmetic problem only – need not be fixed unless extra time is available on the project
2	Minor usability problem – fixing the problem should be given low priority
3	Major usability problem – important to fix, so should be given high priority
4	Usability catastrophe – imperative to fix this before product can be released

Table 4. Five-point rating scale and their descriptions.

The expert review method tends to identify minor problems which do not affect playing the game too much, but are still annoying to the player. Nielsen has claimed that the probability of finding minor problems is

twice the probability of major findings (Nielsen, 1992b). Although the focus of the usability evaluation should be in finding the most critical usability problems, there is also a positive effect on finding minor problems. The minor problems can be the cause of more severe problems and by removing them the developers can increase the quality of the game with minimal effort. It has also been noted that expert review is more efficient in finding those minor problems than other methods (Nielsen, 1992a).

3.4 INSPECTORS

Inspectors play a critical role in how successful expert reviews are. Typically, the inspectors are usability experts who have knowledge about usability principles and the method itself. Quite often it is required that the inspectors have knowledge of the domain for which the system is developed. The inspectors with this knowledge are called *double experts* (Nielsen, 1993). Domain knowledge can also be acquired by recruiting end-users to act as inspectors. It can be assumed that the end-users have very good domain knowledge and the knowledge of tasks that are performed with a system. Følstad et al. (2010) have studied the performance of the work-domain experts and usability experts using group-based expert walkthrough and concluded that the work-domain experts provide valid results although their evaluation may be less thorough.

Another possibility is to let end-users use the system and observe what actually happens instead of asking them to guess what might happen (Nielsen, 1994b) or have domain experts working side-by-side with a usability professional. This technique is especially useful when operating a system requires specialized skills that the inspectors do not have (Chilana, Wobbrock, & Ko, 2010). Nielsen (1993) has noted that even though heuristic evaluation can be performed by people with little or no usability expertise, it is preferable to include usability specialists as inspectors, because the end-users may not know how the redesign of the identified problems might change the system in the future.

Ideally, the expert group will be formed over time and knowledge and expertise is gained by conducting evaluations and educating the inspectors based on the evaluation results. Game development companies are typically small and might not even have a dedicated playability professional, but evaluations are conducted by novice inspectors with very little knowledge and experience of video game evaluations and mobile games in particular. Therefore, well-designed playability heuristics will help the novice inspectors to conduct evaluations more accurately and more efficiently.

Domain expertise in game evaluations concerns the gaming platforms and game genres. Games played on the PC, game consoles, and mobile devices have their own characteristics and the inspectors need to be aware of these differences when they are evaluating games. Especially mobile devices have varying characteristics and without knowing them the inspectors might focus on the wrong aspects in the design or compare them to the conventions in another platform. Knowledge of the game genres is also crucial and the inspectors should know other similar kinds of games that have been published previously because certain playability problems appear repeatedly (Pinelle, Wong, & Stach, 2008b) and design conventions that have matured in specific game genres should be followed in future designs as well. Domain expertise does not require extensive playing time or achieved mastery of a game, but an ability to interpret the gameplay based on their previous knowledge and experiences (Kirschner & Williams, 2013). Skills of playing games are not irrelevant, but including inspectors with various expertise levels will probably bring better insights about playability problems in the design that are relevant for each inspector group. Skills of playing games are needed to progress to advanced levels of a game and to solve more difficult challenges.

It is a common practice that an evaluation is conducted by several inspectors and each inspector will perform the evaluation independently. Nielsen and Molich (1990) observed in their early studies that even if inspectors with the same background evaluated the system, the correlation between the number of usability problems found by individual inspectors was very low. A single inspector will miss most of the usability problems in an interface (Molich & Nielsen, 1990; Nielsen & Molich, 1990; Nielsen, 1992a). The obvious reason for this is that the method is subjective and the inspectors are referring to their previous experience of conducting evaluations and their knowledge of the usability principles and the domain. Better results are achieved when multiple inspectors conduct the evaluation. However, each inspector should conduct the evaluation independently and the evaluation results should be aggregated afterwards. This will ensure unbiased evaluations from each inspector. The inspectors' individual differences in perceptual, psychomotor and cognitive conditions affect the inspectors' performance while interacting with computers (Ling & Salvendy, 2008). The individual differences between the performance of the inspectors is called *evaluator effect* and it affects both user testing and usability inspection methods (Hertzum & Jacobsen, 2001).

Vermeeren, van Kesteren and Bekker (2003) have analyzed several sets of user testing data and concluded that there are five factors which probably affect the interpretation of the data. Most of them are specific to user testing, but one factor, 'inaccuracy of the evaluators,' is a real challenge in expert reviews, as well. Vermeeren et al. describe that the inaccuracy of

the evaluator is related to interpretation and the problems in this process. The inspectors might not remember to report everything that they see, or they may have misinterpreted the instructions that are given to them. Vermeeren et al. (2003) suggest that to overcome this problem, multiple evaluators should carry out the analysis so that multiple views on the problems can be gathered.

Hertzum and Jacobsen (2001) state that the origin of the problems in analytical inspection methods is also in interpretation. They have identified three shortcomings that affect evaluation results: 1) vague goal analysis, 2) vague evaluation process, and 3) vague problem criteria. In an expert review, the goal analysis is related to the inspectors' mental clarity of the evaluation goal and thus, the focus of the evaluation. Although the inspectors agree on the focus of the evaluation, small differences during the evaluation may lead to considerable variability in the evaluation results (Hertzum & Jacobsen, 2001). The vague evaluation process is a critical aspect in an expert review because "the method does not provide a systematic procedure for ensuring that all interface elements are evaluated against all heuristics" (Hertzum & Jacobsen, 2001, p. 435). This has been a problem especially with traditional usability heuristics because "they describe common properties of usable products" (Nielsen, 1993, p. 158). Our playability heuristic set tries to overcome both of these shortcomings by organizing the heuristics into modules which indicate the different focus areas of the evaluation and arrange the heuristics based on their importance in the evaluation.

The third shortcoming 'Vague problem criteria' refers to the threshold when some issue during the evaluation becomes a usability or a playability problem. Hertzum and Jacobsen (2001) give an example from the traditional heuristic list (see Section 4.1) which contains the heuristic 'Match between system and the real world'. This heuristic is particularly difficult for inspectors for the reason that it does not explain what match an inspector should be looking for and how big of a mismatch is allowed before it becomes a usability problem. In our playability heuristic set, we have tried to avoid heuristics which would cause similar difficulties in interpretation. Further, the wording of the heuristics has been defined so that it explains what the concern is when something becomes a problem. Therefore, we have used verbs and other keywords in the descriptions. For example, if a game does not have short-term or long-term goals, it is a playability problem. Probably it is not possible to define heuristics in such a way that they would explicitly indicate when there is a problem in design, but it should be possible to define heuristics which are easy to interpret and apply during the evaluation.

Hertzum and Jacobsen (2001) concluded that the expert review method will always contain the evaluator effect, because the procedure is informal

and leaves plenty of room for the inspector to conduct the evaluation. One solution is to use multiple inspectors. However, we believe that the method can be improved by having a well-defined heuristic set on the right abstraction level which minimizes the possibility of interpretation problems as suggested by Cronholm and Bruno (2008).

3.5 PLAYTESTING IN GAME DEVELOPMENT

User testing is probably the most common evaluation method to assess the usability of a product. In game evaluations, this method has been the primary method for years.

When looking into game design literature and how it talks about playtesting, there are a couple of things that come up regularly. It seems that game designers have a somehow different view of what user testing (or playtesting) actually means compared to HCI literature (e.g. Rubin, 1994; Dumas & Redish, 1999; Dumas & Loring, 2008). The first notion is that game designers like to see usability testing as separate from playtesting. In their mindset, usability testing mainly concerns interaction with the game user interface and whether the system is intuitive and easy to use (Fullerton et al., 2004; Schell, 2008). The objective of playtesting is to gain insight into whether the game meets player experience goals (Fullerton et al., 2004; Schell, 2008). In other words, whether the game is balanced and fun to play. Although game design literature separates usability testing of productivity software from playtesting, the procedure of conducting playtesting with external players is almost identical. It includes an introduction part, an explanation of the test procedure and the player's role in the test, an actual test session and a post-test interview and rewarding of the participants (Fullerton et al., 2004).

The second notion is that game designers seem to be very cautious when a game is shown to the target player population and they prefer playtesting the game either internally or with persons they know well and trust (Rouse, 2001; Fullerton et al., 2004; Schell, 2008). It is a common practice that a game is first tested by the members of the development team. These tests are called playtesting even though their main objective is to enthuse members about the project and increase their understanding of how the game might be improved (Rouse, 2001). Fullerton et al. (2004) say that it is an internal design review rather than a playtesting session.

The second group of testers is usually quality assurance (QA) play testers and other game designers who are not a part of the project team. These testers are usually considered hard-core gamers and they are skilled players who might get used to certain problems and may not notice problematic issues for other players. There is also a risk that if too much weight is put on their comments, the game is too hard for average players

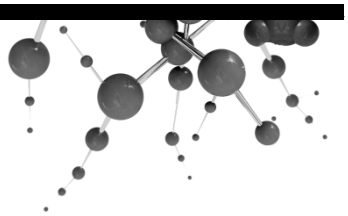
(Schell, 2008). The third group of testers is first-impression testers and non-gamers. They will provide feedback as players who have not seen the game before or do not have game design knowledge.

Playtesting is applied throughout the development process, but testing is mainly done internally and real players are involved in later stages. In some cases, potential players are invited to an early testing to get enough players for stress testing, but in this case, players are merely resources for QA testing, and the test is not focused on them (Mulligan & Patrovsky, 2003). Rouse (2001) points out that in early playtesting only experienced professionals should be used as play testers because only they can overlook problems and incompleteness that the game design has. Design literature also recommends using friends and other confidants as play testers of a game when it is too early to be shown to external players. Interestingly, the same authors also warn that these confidants are unreliable and their feedback may be biased for various reasons (Rouse, 2001; Fullerton et al., 2004; Schell, 2008). Fullerton et al. (2004) and Schell (2008) say that getting criticism and negative feedback from strangers can be hard for game developers which is probably one of the reasons why external players are not used very often. However, productivity software development has the same problem, but there it has been solved so that in user testing there are separate moderators who are not a part of the development team, which makes it possible to give and receive negative feedback neutrally.

External players are involved once the user interface and the controls need to be tested. However, it is still possible that as the test focuses on controls, only a limited gameworld, level or situation is provided to the testers (Rouse, 2001). This playtesting resembles an exploratory test of productivity software (Rubin, 1994) where only a fraction of the product is tested. The game is usually shown to real players when it is almost ready to be shipped. This testing phase is usually called beta testing (Mulligan & Patrovsky, 2003). However, it is usually too late to make any big changes in gameplay (Rouse, 2001) because the game needs to be "feature-complete" when beta-testing starts.

One of the specialties of playtesting, which does not exist in productivity software testing, is the testing for game balance. Balancing means that some settings or attributes of game items or characters are tweaked. Playtesting is then used to see whether those changes result in the right amount of challenge for the players (Rouse, 2001). The critical issue here is that the whole game should be balanced equally, and the game designers need to ensure that balancing works for the entire game. Therefore, effective balancing can be done when the most of the game is ready and the game content is complete (Rouse, 2001). Players from the target player population can give valuable feedback on how difficult the game is.

Based on the game design literature review of playtesting it seems that the expert review method would be useful for game designers in many respects. Playability experts are used to seeing incomplete versions of games and they can ignore issues which are not relevant in the evaluation round. Playability experts are often skilled players who have extensive knowledge of different games, but as outsiders they can provide a fresh view to the game development project. The most important characteristic of playability experts is that they can give feedback which brings valuable insights for game designers. Their comments are not biased and they can analyze their own behavior while playing a game.



4 Heuristics to Aid Inspection

In this section, we will take a look at general usability heuristics that have been developed to assist in usability inspections. Then we will explore domain-specific heuristics that other researchers have developed to evaluate the playability of video games.

4.1 General Usability Heuristics

The general usability heuristics have been developed and presented by Nielsen and Molich (1990). The original heuristics were derived empirically from the factor analysis of 249 usability problems (Nielsen, 1994a) and it contained nine heuristics. Later on the heuristics were expanded to include a "*Help and Documentation*" principle which was added in 1991. The heuristics contain ten general principles on user interface design. The wording and terms used in the heuristics have been changed during the years in different publications (Nielsen & Molich, 1990; Nielsen, 1994a). The following list of heuristics and their descriptions is available on the Nielsen Norman Group web site[4].

1. **Visibility of system status**

The system should always keep users informed about what is going on, through appropriate feedback within a reasonable time.

2. **Match between system and the real world**

The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, make information appear in a natural and logical order.

[4] http://www.nngroup.com/articles/ten-usability-heuristics/

3. User control and freedom

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. Error prevention

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. Recognition rather than recall

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for the use of the system should be visible or easily retrievable whenever appropriate.

7. Flexibility and efficiency of use

Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. Aesthetic and minimalist design

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no codes), indicate the problem precisely, and suggest a solution constructively.

10. Help and documentation

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

As it can be seen, usability heuristics mainly concentrate on the user interface of the product. They do not consider the application domain or content that the user interacts with the product. The heuristics were also developed on purpose to be fairly broad so that they can be applied to practically any type of user interface design (Molich & Nielsen, 1990).

4.2 DOMAIN-SPECIFIC HEURISTICS FOR DIFFERENT DOMAINS

Many inspectors have found that the general usability heuristics do not always meet the specific needs of different product categories, and they require supplemental guidelines or some re-interpretation to make sense (Zazelenchuk, 2006). Nielsen has also noted that it is possible to develop heuristics that apply to a specific class of products and supplement the general heuristics (Nielsen, 1994b, p. 29). The purpose of the domain-specific heuristics is to improve their reliability and 'goodness of fit' for new products and domains (Zazelenchuk, 2006).

Over the years, several heuristic sets (Table 5) have been defined for different domains. Most of them are targeted for certain product categories, but there are also more generic ones to cover things like accessibility (Paddison & Englefield, 2004), security (Jaferian, Hawkey, Sotirakopoulos, Velez-Rojas, & Beznosov, 2014) and Virtual Worlds (Munoz, Barcelos, & Chalegre, 2011). Further, there are heuristic sets for certain kinds of devices, such as medical devices (Zhang, Johnson, Patel, Paige, & Kubose, 2003) or touch-screen devices (Inostroza, Rusu, Roncagliolo, Jimenez, & Rusu, 2012). Many of the domain-specific heuristic sets are actually modified versions of the general usability heuristics. The authors have made only small changes to the wording and perhaps added a few new heuristics which deal with the characteristics of the domain. This limits their usefulness because there is not much added value compared to the general usability heuristics.

Social TV heuristics (Geerts & De Grooff, 2009) focus on social aspects of TV viewers and watching TV programs. This heuristic set is a particularly interesting one because it points out similar things that are included in our multi-player heuristics (Section 5.3). The heuristic set includes aspects such as communication between users and presence and awareness of the users, which are essential for social interaction both in collocated and remote occasions. The interaction can happen between users and the system and the users should be able to control what information is shared between the users and what are the actions the users can take when watching TV.

The heuristic set should not only cover positive aspects of the domain, but it should analyze the domain from different perspectives. The social TV heuristic set is a good example of this as it includes a heuristic for privacy of the user and possible interference that new features possibly cause for the basic activity of watching a TV program. These are aspects that inspectors should pay attention to as well when they evaluate social TV applications and systems.

Domain	Reference	Heuristics	Originality
Accessibility	(Paddison & Englefield, 2004)	9	Novel
Ambient Displays	(Mankoff et al., 2003)	7	Modified
Computer Assisted Assessment	(Sim, Read, & Holifield, 2008; Farrell & Farrell, 2012)	11	Modified
E-learning applications	(Benson et al., 2002)	15	Modified
	(Evans & Sabry, 2003)	9	Novel
	(Ardito et al., 2004)	18	Novel
	(McKay & Kölling, 2012)	13	Modified
Groupware Systems	(Baker, Greenberg, & Gutwin, 2001; Drury, 2001)	8	Novel
Information Visualization	(Zuk et al., 2006)	13	Novel
Interactive Web Sites	(Petrie & Power, 2012)	21	Novel
Medical Devices	(Zhang et al., 2003)	14	Modified
Mobile Map Applications	(Kuparinen, Silvennoinen, & Isomäki, 2013)	10	Modified
Notification System	(Berry, 2003)	8	Modified
Security	(Jaferian et al., 2014)	7	Novel
Social TV	(Geerts & De Grooff, 2009)	12	Novel
Touchscreen-based Mobile Devices	(Inostroza et al., 2012)	11	Modified
Virtual Reality Applications /	(Sutcliffe & Gault, 2004)	12	Novel
Virtual Worlds	(Munoz et al., 2011)	16	Modified

Table 5. Domain-specific heuristics for different domains.

4.3 DOMAIN-SPECIFIC HEURISTICS FOR VIDEO GAMES

Game researchers have questioned the applicability of traditional usability heuristics for game evaluations. In their study, Johnson and Wiles (2003) show that the traditional usability heuristics conflict in game evaluations to achieve a good player experience. Federoff (2002) analyzed general usability heuristics and found out that even though some heuristics are applicable for game evaluations, especially when evaluating the interface of the game, the heuristics fail in the ability to address gameplay issues. She believes that game heuristics should encompass the three major areas: game interface, game mechanics, and game play, which were identified by Clanton (1998). Hence, game researchers have started to develop heuristics which would include both usability and gameplay issues, to assist game developers in discovering playability problems in game design (Nokia, 2006).

One of the first studies exploring the playability of video games was conducted by Malone (1980) when he studied video games to find out what makes them enjoyable. In the study, three factors were identified to be major components: challenge, fantasy, and curiosity. In fact, these three factors are broader themes that include various aspects which are essential for enjoyable experiences in games. Challenge is related to difficulty in using skills and achieving goals. Curiosity is a motivation to learn the game and use of skills to accomplish goals. Fantasy is related to narrative aspects or the theme of the game and how skills are used in the game. Although the study was conducted a long time before the expert review method even existed, these factors could be considered as an initial list of playability heuristics.

After Malone's study, there have been isolated studies about fun and HCI aspects in games (e.g. Cherny et al., 1997; Clanton, 1998). Garris et al. (2002) defined six game dimensions that describe game characteristics based on a literature review: fantasy, rules/goals, sensory stimuli, challenge, mystery, and control. Clanton (1998) defined 15 design principles that can be considered important for game design, but they were not developed into the form of playability heuristics. These principles are:

- Establish a Quest
- Provide a gentle on-ramp
- When players select a difficulty, they accept it
- Let each player progress at their own rate
- Spread clues, tools and obstacles out but not too much
- Avoid lengthy dead ends
- Pressure can be fun
- Give hints not answers
- Avoid linear, monotonous pacing
- Reward gameplay with media

- Confusion is not fun
- Frustration can be fun
- Trial and error is not fun
- It's fun to be known
- Make a great game and the players will master its complexity

Clanton had collected design principles from different game genres: action games, strategy games, adventure games, puzzle games and role-playing games. Cross-checking multiple genres is one of the fundamental issues in defining playability heuristics. In order to define a good set of playability heuristics, they should be applicable for several game genres.

The systematic development of playability heuristics started in 2002, when Federoff (2002) studied a game company to find out how to assist a game development project to produce successful games more consistently and incorporate fun into the design and possibly improve customer satisfaction. Based on a literature review of five game designers and researchers, and interviews of personnel in the company, she defined 40 heuristics related to the game interface, game mechanics, gameplay and game story. Federoff did not validate the heuristics so it is unknown how effective they are in actual game evaluations, or at least according to my best knowledge such results have not been published. Desurvire et al. (2004) continued Federoff's work and published Heuristic Evaluation for Playability (HEP) which was a modified list of the heuristics. The list was only slightly changed and many of the heuristics were the same as in Federoff's heuristic set. Desurvire et al validated the heuristics by comparing evaluation results to user testing results. They concluded that the heuristics successfully identified playability problems in a game prototype.

At the same time Fabricatore et al. (2002) published a qualitative model of player preferences in action video games that would determine the playability in these games. The model is a hierarchical structure of categories of concepts (Figure 12) considered important by the players. Three main issues are Entity (Avatar), Scenario (Gameworld) and the hierarchy of goals. The issues are presented as design recommendations for action games, but they could be converted into heuristics as well.

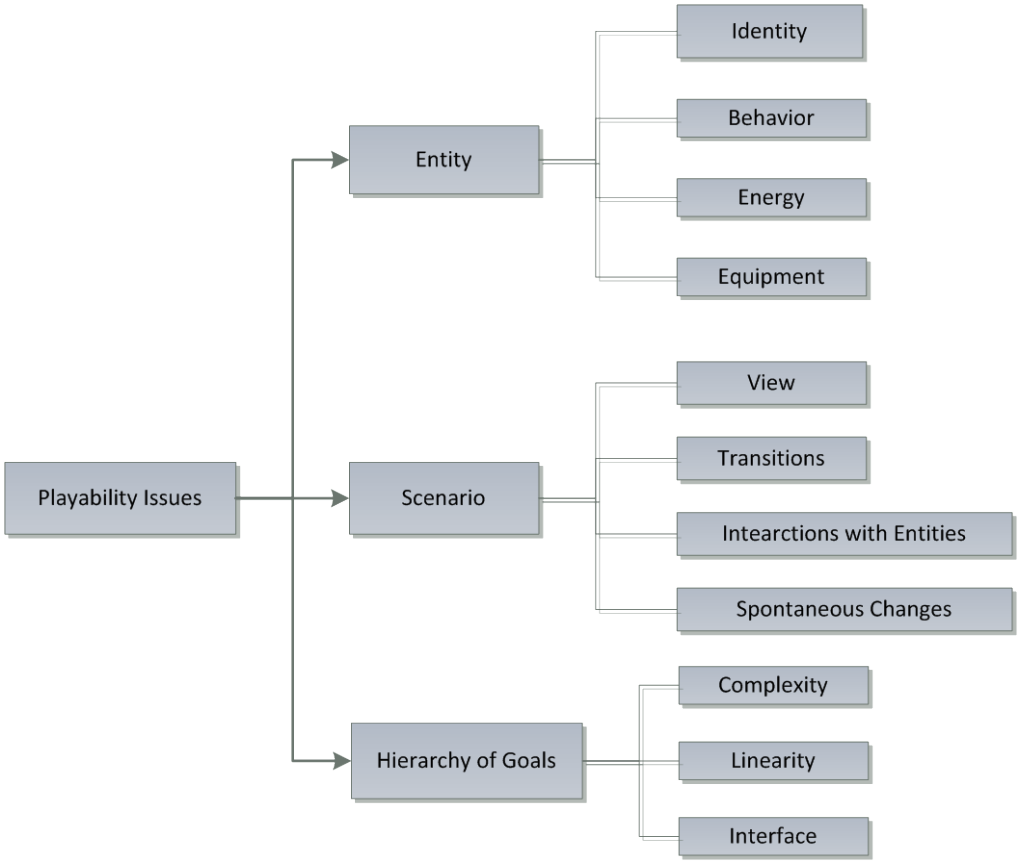


Figure 12 Hierarchy of playability issues in action video games (Fabricatore et al., 2002).

The mobility aspects of playability were included in the research agenda in 2005 when mobile games were an emerging trend in gaming. We started our research to define playability heuristics which included single player games and mobility aspects (Korhonen & Koivisto, 2006; Nokia, 2006). Karvonen (2005) explored the playability of mobile games in his master's thesis and introduced a framework that would highlight a player-centric view of playability. The framework was developed based on a literature review and game critiques from different web sites that have specialized in reviewing games. The games were selected based on the average scores of the reviews that can be found on the Metacritic website (Karvonen, 2005). The framework is independent of game genres and platforms in order to provide a commonly used framework for evaluating playability. Based on the used method and data, the framework also includes aspects that are outside the scope of playability and more related to the overall player experience. The framework has not been used with an expert review, but it was used to define questions for a structured interview.

The next logical step in playability heuristic set development was to expand the current heuristic sets to include multi-player games. We introduced multi-player heuristics (Korhonen & Koivisto, 2007) in 2007. These heuristics covered communication, presence and interaction

demands of the players to be able to play together. Another heuristic set for multi-player games was presented by Pinelle et al. (2009).

After the initial introduction of playability heuristics, other researchers started to complement heuristic sets. Schaffer (2007) introduced heuristics for the usability in games in a white paper. According to him, previously published heuristics lacked concrete examples. This makes them less clear for practitioners. Schaffer's heuristics are based on a literature review of earlier works and on his own expertise in the field (Schaffer, 2007). Pinelle et al. (2008a) focused on usability issues and published 10 game usability heuristics. They were based on an analysis of PC game critiques from a well-known gaming website. The analysis included six major genres and 108 games. The heuristics were developed based on common problem categories presented in those game critiques. The heuristics focused on game-specific usability issues related to understanding, learning, and controlling the game (Pinelle et al., 2008a). The gameplay issues are not covered in the heuristic set. Papaloukas, Patriarcheas and Xenos (2009) explored heuristics for specific hardware configurations or social games. Desurvire and Wiberg (2009) have presented PLAY heuristics, which are based on the earlier HEP heuristics. The PLAY heuristics feature 19 top level headings, each containing one to six heuristics (50 heuristics in total). The PLAY heuristics are aimed toward specific game genres such as real-time strategy games, first person shooters and action adventure games.

Genre-specific heuristics have been developed by several researchers. Köffel and Haller (2008) presented a set of heuristics (10) for advanced electronic tabletop games. Their list was a synthesis of earlier heuristic sets. Paavilainen (2010) has proposed new heuristics for social games after reviewing the existing playability heuristics and concluding that they do not meet the specific needs of social games. Recently, Hara and Ovaska (2014) have defined heuristics for the interaction design of motion-controller games.

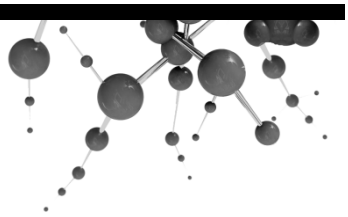
Educational games have received special attention from researchers and several heuristic sets have been developed for these games. Thomas et al. (2003) and Kiili (2005) were the first ones to present their heuristics or design principles for educational games. The principles presented by Kiili are particularly interesting because they cover the basic games design principles of a player's skills, challenge, control and feedback, but the perspective is to support the flow experience (Kiili, 2005). Later on other researchers have presented their own or modified versions of heuristics for educational games (Hinske et al., 2008; Mohamed & Jaafar, 2012).

Table 6 presents a summary of the playability heuristic sets that have been published for video games. It seems that the development of playability heuristics for games is not yet over and there will be new emerging needs to fine-tune the heuristics for specific purposes.

Author(s)	Year	No. of Heuristics	Applicable	Originality
Fabricatore et al.	2002	11	Genre-specific	Novel
Federoff	2002	40	Generic	Novel
Thomas et al.	2003	5	Educational	Modified
Desurvire et al.	2004	43	Generic	Modified
Karvonen	2005	30	Generic + Mobile	Novel
Kiili	2005	22	Educations	Novel
Korhonen & Koivisto	2006	29	Generic + Mobile	Novel
Korhonen & Koivisto	2007	8	Multiplayer	Novel
Schaffer	2007	21	Generic	Modified
Köffel & Haller	2008	11	Genre-specific	Modified
Korhonen et al.	2008	4	Context-Aware	Novel
Pinelle et al.	2008	12	Generic	Novel
Desurvire & Wiberg	2009	50	Generic	Modified
Hinske et al.	2009	33	Educational	Novel
Papaloukas et al.	2009	10	Generic	Modified
Pinelle et al.	2009	10	Multiplayer	Modified
Paavilainen	2010	10	Social Games	Novel
Mohamed & Jaafar	2012	42	Educational	Modified
Hara & Ovaska	2014	13	Genre-specific	Novel

Table 6. A summary of playability heuristic sets developed for game evaluations.

There are also other guidelines to help game developers in the design process and to make more engaging and usable games for players (e.g. Falstein & Barrywood, 2006; Snow, 2007). In addition to articles and websites, edited books have also been published on the topic (e.g. Isbister & Schaffer, 2008; Bernhaupt, 2010), which cover various methods for evaluating usability, playability and user experience in general.



5 Playability Heuristics for Mobile Games

In this section, we will finally take a look at the heuristics that have been developed during the research work to help in evaluating the playability of video games. The playability heuristic set contains 47 heuristics organized into five modules (Figure 13). The inspectors can include or exclude each module depending on the needs of the evaluation. Two core modules, *Gameplay* and *Game Usability*, are common to all games. The Game Usability module covers game controls and the interface through which the player interacts with the game. The Gameplay module contains heuristics that cover the game mechanics and gameplay issues in the game content. The *Multi-player* module covers issues which are relevant for the social interaction of the players. The *Mobility* module contains heuristics that are specific for mobile games. And finally, the *Context-Aware* module deals with issues related to contextual factors in the mobile game design.

The heuristics were originally described in one of Nokia's technical reports (Nokia, 2006). In the following, each heuristic is presented in parts: a title, a short summary, background and similar heuristics. The title and the short summary can be used during the evaluation to support the inspectors' task (Appendix A presents a consolidated list of heuristics that inspectors can use directly). They will help in interpreting a playability problem and assigning a violated heuristic. Background information is used in this dissertation to give a better description of the heuristic with references to relevant literature. There is also a list of similar heuristics if such heuristics exist in other heuristic sets. Some of our heuristics have been influenced by the work of earlier researchers. Correspondingly, some heuristics presented in this dissertation have been adapted to newer heuristics sets.

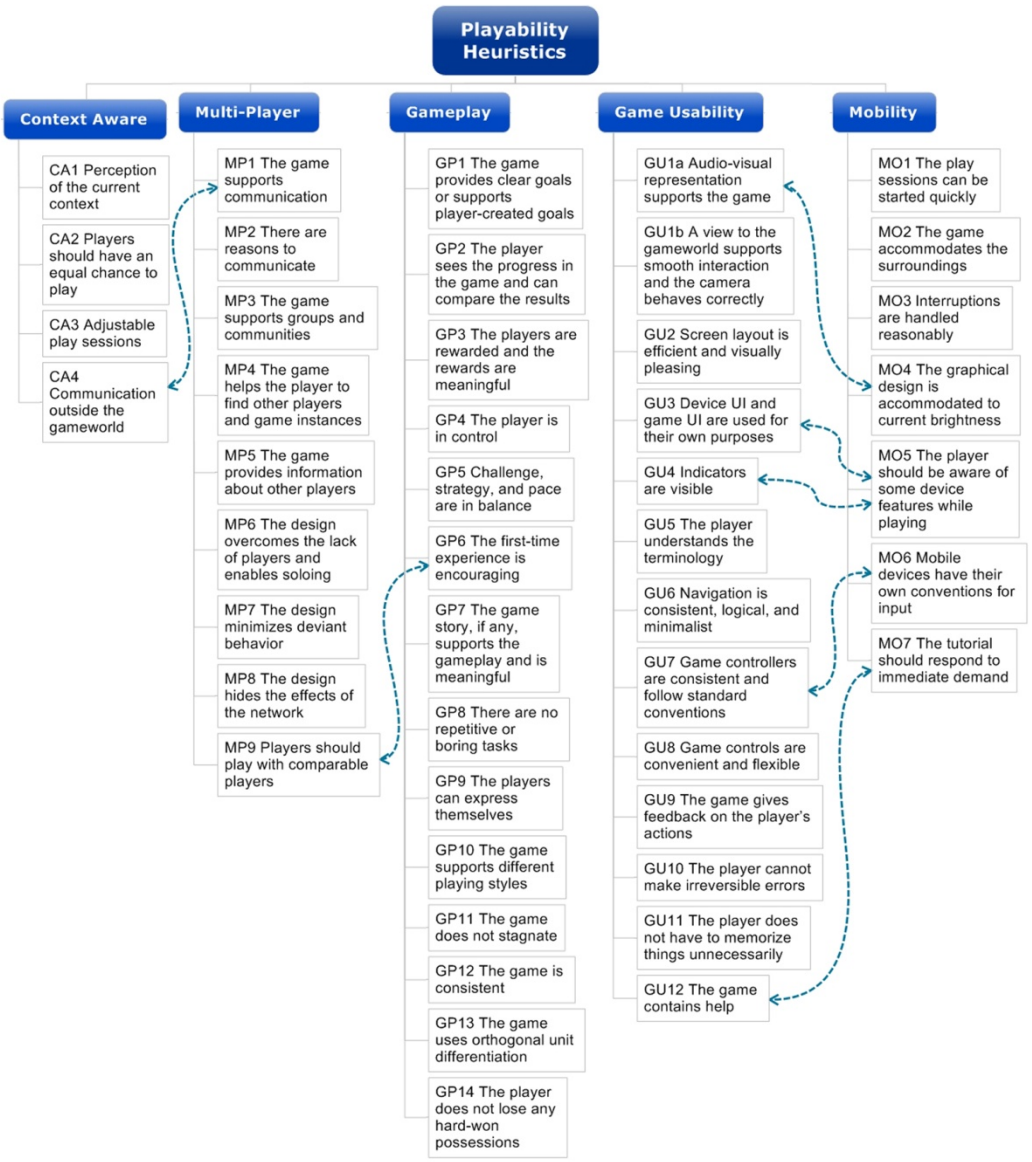


Figure 13. Playability heuristic set and relationships of individual heuristics between modules.

5.1 GAME USABILITY HEURISTICS

Game usability heuristics cover aspects which are similar to traditional usability issues of productivity software, but they are transformed to cover game related aspects specifically. Moreover, there are other game-related issues that need to be covered by the heuristics.

The heuristics have been published in Paper I.

Based on our playability evaluations of social games (e.g. Paavilainen et al., 2012) and a game design literature review, I made one modification to the heuristic set that we have used for years in playability evaluations. In these studies, we have encountered a problem with the viewing angle to the gameworld multiple times. This has been the case especially with social games that are implemented using technology that is not flexible enough to support a proper 3D view. The gameworld is usually presented as a 3D world, but the viewing angle cannot be rotated or panned which means that some gameworld items will be hidden behind other items. Previously these problems have been categorized under the first game usability heuristic GU1 "*Audio-visual representation supports the game*". However, the heuristic does not describe the problem accurately because it focuses more on the aural and aesthetic part of the game and not on the viewing angle to the gameworld. Therefore, a new heuristic called GU1b "*A view to the gameworld feels natural and the camera behaves correctly*" has been introduced to cover playability problems related to views to the gameworld and camera behavior. Furthermore, the identifier of the original heuristic GU1 has been changed to GU1a to indicate the split.

GU1a: Audio-visual representation supports the game

Game graphics should support gameplay and story and be informative for the player. In addition, the graphical look and feel should be consistent throughout the game. Audio can be used to evoke emotions and increase immersion. A good sound environment in the game supports a positive gaming experience. The graphics or audio should not prevent the player from performing actions or make them unnecessarily difficult.

The reason for including this particular heuristic is quite straightforward as music and graphics are strong determinants of the aesthetics of a game (Schell, 2008). Audiovisual style and appearance can have a significant influence on player experience (Järvinen, 2002a). Graphics are often treated as a decorative attribute that is meant to increase eye-candy of the game, but it has a key role in creating sensory immersion of the players as well (Ermi & Mäyrä, 2005). Sounds and graphics can be used to create sensory curiosity and enhance the fantasy of the game (Malone, 1982). From the marketing point of view, many games are sold because of their visual quality and players expect that games are visually appealing.

From the evaluation point of view, the inspectors should pay attention to the game's graphic style and whether it supports the game mechanics (Egenfeldt-Nielsen et al., 2008) and is convenient for the game concept (Järvinen, 2002a). Graphic style should be tied to other components of the game (including a possible story element) and be informative for a player. Consistency of the graphical style and appearance is also an important factor for playability (Adams & Rollings, 2007).

There are three kinds of graphic styles (*photorealism*, *caricaturism* and *abstractionism*) that are frequently used in games and the proper style depends on the game concept (Järvinen, 2002a). Photorealism simulates environments and characters as closely as possible to reality. Caricaturism presents a simplified version of the characters and the environment by emphasizing the most characteristic features. Pure abstractionism is rare because most games simulate the environment and present forms in a caricaturized rather than a completely abstract way (Järvinen, 2002a).

A proper audio style is much more difficult to evaluate because there are numerous possibilities to use audio in games. In addition, it is easy to underestimate the importance of audio in games because graphical elements are dominating the output channel. However, a proper use of audio can reinforce the premise, evoke emotions and inform the player when something important happens in the game (Barry, 2005). In general, it could be said that a properly implemented sound environment amplifies a positive game experience of the player.

There are several audio elements (sound effects, music and dialog) which are commonly used in games (Barry, 2005). Adams and Rollings (2007) have supplemented the list with ambient sounds and voice-over narration which will create a complete audio environment in the game.

Sound effects are the most common use of sound in game design and they are used both in the shell menu and in the gameworld. These sounds correspond to actions and effects in the gameworld and give feedback for the player's actions (Barry, 2005; Adams & Rollings, 2007). Sounds should be harmonized between the shell menu and the gameworld to provide a seamless sound environment.

Ambient sounds are not used very often in other software products, but in games they are present. Ambient sounds work together with the main view to the gameworld. The main view gives visual feedback about where the player is at the moment. Ambient sounds deliver aural feedback about the player's current location. Adams and Rollings give a warning that ambient sounds should not be overused especially in situations where mental challenges are present (Adams & Rollings, 2007).

Music is a strong audio element and it is often used for creating a certain mood for the player or the atmosphere of the game or reinforce the theme (Barry, 2005; Schell, 2008). Music sends a strong cultural messages and it should be harmonized with the gameworld and the gameplay (Adams & Rollings, 2007). One critical issue with music in games is that players will hear it for several hours (Adams & Rollings, 2007). There is a risk that the music becomes annoying after a while. Some games allow the players to use their favorite music instead of original songs in the game. The music can also be used for defining what players will do. It is a sort of indirect controlling mechanism (Schell, 2008).

Dialog and voice-over narration are useful audio elements with games that have strong storylines (Adams & Rollings, 2007). Although they provide an excellent form of feedback in many games, there are also challenges that are related to recorded dialogs. A player can easily get tired of hearing the same words repeated over and over again. Therefore, there should be multiple variants for recoded dialogs which occur repeatedly and they should be mixed up at random (Adams & Rollings, 2007). Another challenge is that voice acting must be of good quality and the language that is used in recordings must be harmonized with the theme of the game. Usually voice prompts are used to give a game character a voice (Barry, 2005).

Similar heuristics:

- "Use sound to provide meaningful feedback." (Federoff, 2002)
- "Art should speak to its function." (Federoff, 2002)
- "Sounds from the game provide meaningful feedback or stir a particular emotion." (Desurvire et al., 2004)
- "Art should be recognizable to player, and speak to its function." (Desurvire et al., 2004)
- "Sounds (how the game sounds): the quality of sound effects, music, and dialogs, suitability for the game, provide configuration possibilities [translated from Finnish]" (Karvonen, 2005)
- "It should be clear what's happening in the game. Players should understand and be able to identify game elements like the Avatar and Enemies, Obstacles, Power-ups." (Schaffer, 2007)
- "Art is recognizable to the player and speaks to its function." (Desurvire & Wiberg, 2009)
- "Provide appropriate audio/visual/visceral feedback (music, sound effects, controller vibration)." (Desurvire & Wiberg, 2009)
- "Visual representations, such as maps, icons, and avatars, are frequently used to convey information about the status of the game. Visual Representations should be designed in an easy to interpret way, and so that users can differentiate important elements from irrelevant elements." (Papaloukas et al., 2009)

GU1b: A view to the gameworld supports smooth interaction and the camera behaves correctly

The view to the gameworld defines how the player perceives that gameworld and how well the player can immerse in the game events. The gameworld can be presented either in a 2D or a 3D view and there is no single best option of perspective, but it depends on the type of game and the game concept. It is more important to notice how the perspective influences the player's ability to interact in the gameworld.

One of the biggest things affecting the playability of a game is the view to the gameworld or more specifically the perspective from which the gameworld is viewed by a player. Basically all games (except text-based games) include a view to the gameworld (Egenfeldt-Nielsen et al., 2008). The perspective will influence greatly how players perceive the gameworld and how close the player can get to game avatars and objects (Egenfeldt-Nielsen et al., 2008). This will affect how well the player can get immersed in the game events.

The gameworld can be presented either in a 2D or a 3D view. There is no single best option of perspective, but it depends on the type of game and the game concept (Clanton, 2000; Adams & Rollings, 2007; Egenfeldt-Nielsen et al., 2008; Schell, 2008). Egenfeldt-Nielsen et al. (2008) have concluded that any perspective is able to provide a positive player experience. Some genres and subgenres constantly adhere to one or very few perspectives, but it is important to evaluate whether the perspective is appropriate for the game concept. A 3D camera rotation or angle may look nice and cool, but it may also influence the player's ability to interact in the gameworld and make the game unnecessarily difficult to play.

In 2D views, the gameworld can be presented in a single static screen that shows a complete gameworld to the player or in a view that utilizes either side-scrolling or top-scrolling (Adams & Rollings, 2007; Egenfeldt-Nielsen et al., 2008). In side-scrolling, the view to the gameworld moves forward and backward and the player sees the gameworld from the side as the camera tracks the game avatar. In top-scrolling, the view to the gameworld is from the top and the gameworld moves up or down. Static screens can also be used as a scenery view in which the gameworld is shown as tiles and the player moves between sceneries.

3D views are commonly used in many commercial games because they provide a more realistic view to the gameworld and allow more freedom for players to move in the gameworld. Table 7 shows typical 3D views that are commonly used in game designs: first-person perspective, third-person perspective, and aerial perspective (Järvinen, 2002a; Barry, 2005;

Egenfeldt-Nielsen et al., 2008). Further, the aerial perspective can be divided into subcategories depending on the camera view.

View	Description
First-person perspective	The player can usually only see a small fraction of the game avatar and the main focus is on the terrain ahead (Adams & Rollings, 2007). This view is very efficient in creating immersion as the player and the game avatar see the gameworld in a similar way.
Third-person perspective	The player sees the game avatar from behind and the camera normally follows the avatar at a fixed distance and slightly above (Adams & Rollings, 2007).
Aerial perspective	The player sees a large part of the gameworld which allows the player to control numerous game avatars at once. In isometric perspective, the player looks at the gameworld from an angle. In top-down perspective, the gameworld is shown directly overhead and the camera is pointing straight down (Adams & Rollings, 2007).

Table 7. Typical 3D views to the gameworld.

The first-person perspective is often assumed to connect a player closely to their game character. However, players usually assume and use their own identity, if it is not reinforced by the game. In addition, navigating through game spaces often requires that the players will get familiar with the dimensions of the game and how the game character fits into the space (Barry, 2005). An over-the–shoulder view makes it easier for the player to connect to the game character. However, there are problems with an accurate direction that the game character is pointing (Barry, 2005).

Third-person perspective is commonly used in video games and it lets the player see the avatar (Adams & Rollings, 2007) or a number of objects (Egenfeldt-Nielsen et al., 2008). The camera follows the avatar and is located behind and slightly above the avatar, thus allowing the player to see both the avatar and the environment. The negative aspect of this view is that the player does not see the avatar's front or side. In addition, when the avatar moves quickly, the camera must sweep around quickly in order to remain behind the avatar. This can cause motion sickness in the player (Adams & Rollings, 2007, p. 244).

Aerial perspective gives priority to the gameworld and it is commonly used in games which allow the player to see any part of the gameworld. Normally the player would see the gameworld directly from above, but as this is not a very interesting view, buildings and other objects are drawn

slightly tilted. Another option is to use an isometric perspective. With this view, the player can feel himself little remote from the action and less attached to the outcomes (Adams & Rollings, 2007, p. 246).

Egenfeldt et al. (2008, p. 107) notes that video games seem to work equally well in both the first- and the third person perspective and proving positive player experience. Perspective has an important role in shaping the player's perception of the game world, avatars and game items (Egenfeldt-Nielsen et al., 2008, p. 113).

Similar heuristics:

- "Implementation quality: Technical feasibility, correctness, loading times, camera angles and control of the camera." [Translated from Finnish] (Karvonen, 2005)
- "Provide unobstructed views that are appropriate for the user's current actions." (Pinelle et al., 2008a)

GU2: Screen layout is efficient and visually pleasing

The layout should present all necessary information to the player, but on the other hand, if the screen is filled with all kinds of information, it starts to look crowded. It is important that the player finds the navigation controls and they should not be mixed with the information that needs to be visible on the screen.

Figure 14. Shell Menu (A screenshot from Lord of the Rings Online, Turbine, 2008).

Usually a game includes two user interfaces, which both have substantially different roles and require different kinds of user interface designs. After launching the game, a player usually enters a shell menu. The shell menu is used to provide access to features which are not part of the gameworld, but are needed for configuring and managing the game (Figure 14). The shell menu contains different kinds of settings and configuration screens for video and audio settings, and a selection of a game mode (e.g. single player vs. multi-player mode). In addition, it contains user interface elements for managing saved game instances, selecting a game avatar and

difficulty settings, and commands for managing a play session. Johnson and Wiles (2003) have noted that shell menu designs are often poorly or non-intuitively organized leading to confusion and errors of the players. During the gameplay, a subset of shell menu functions is typically available in a pause menu.

During the gameplay, the user interface consists of two layers. The game is played through a main view that presents the gameworld to a player. The main view also contains on-screen user interface elements which are used to give input commands (Barry, 2005) and output elements which the player needs to see while playing the game.

Adams and Rollings (2007, p. 235) have studied more than 2000 screen layouts from different kinds of games and identified nine designs which are frequently used in main view designs. Figure 15 shows how the main view is usually divided between the view to the gameworld (white areas) and which parts of the screen are usually reserved for feedback controls and other on-screen controls (grey areas). In actual game designs, the on-screen control area is often transparent so that it does not limit the view, but the player is able to see the gameworld behind and between the controls. In addition, game designers use pop-up controls which become visible in a certain situation or upon request. Pop-up controls can appear anywhere on the screen.

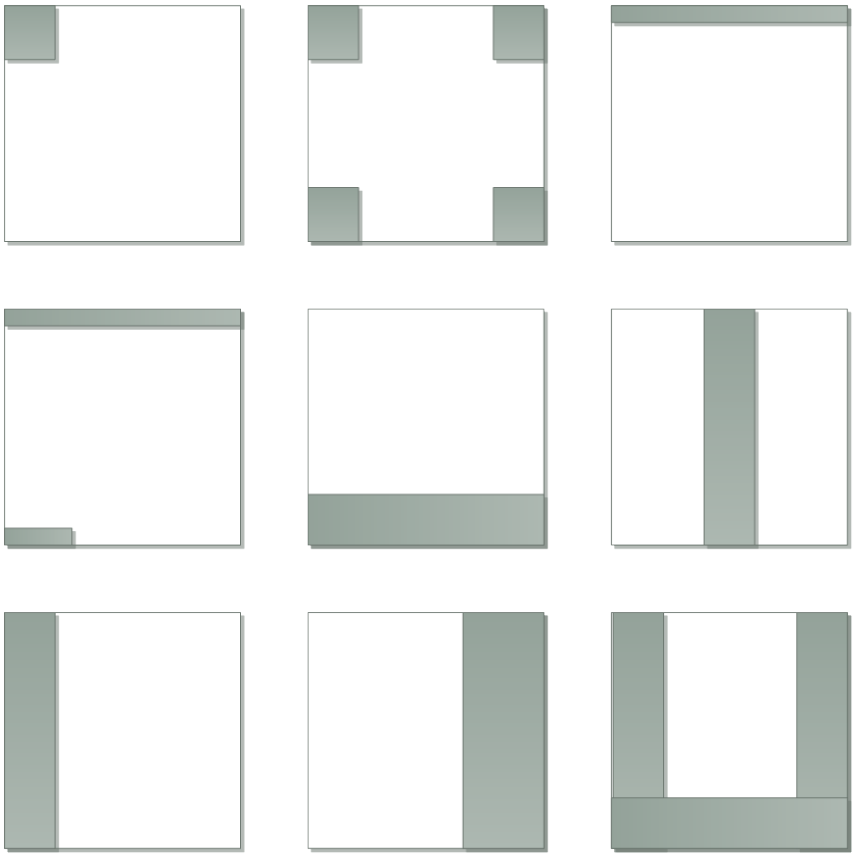


Figure 15 Common views to the gameworld based on Adams and Rollings (2007).

Designing an efficient and good-looking layout is not an easy task. The player should be able to see the necessary information in a single glance rather than have to look around from different screens (Adams & Rollings, 2007). The player must be able to constantly make decisions based on the visible data because some information is hidden deliberately. User interface design for games should follow the same principles that apply to screen layout design in general (Fox, 2005). The screen can get easily crowded with information that might be relevant for the player. One of the important things is that navigation controls are distinguishable and they should not be mixed with the information that needs to be visible on the screen.

Similar heuristics:

- "Minimize the menu layers of an interface." (Federoff, 2002)
- "Make the menu layers well-organized and minimalist to the extent the menu options are intuitive." (Desurvire et al., 2004)
- "The Player should experience the menu as a part of the game." (Desurvire et al., 2004)
- "Don't bury frequently used information." (Schaffer, 2007)
- "Avoid large blocks of text." (Schaffer, 2007)
- "Don't display irrelevant information." (Schaffer, 2007)
- "Screen layout is efficient, integrated, and visually pleasing." (Desurvire & Wiberg, 2009)
- "The players experience the user interface/HUD as a part of the game." (Desurvire & Wiberg, 2009)

GU3: Device UI and game UI are used for their own purposes

It should always be noticeable whether the player is dealing with the game user interface or device functions. The game interface should not use the device's user interface widgets in the game interface, because it breaks the immersion. The most impressive immersion is achieved when the game uses full-screen mode hiding other features.

Video games are frequently played on devices which are capable of multitasking and the game is an application among others. The purpose of the visual design of the game is to immerse the players in the game and separate the game from other applications. Preferably a full-screen mode should be used which hides other features and applications of the device. In addition, a unified theme between the user interface design and the game concept will also reinforce the intended experience of the game (Schell, 2008).

It is possible to use similar kinds of user interface widgets that are used in non-game applications to control processes which are too complex to be

managed with game controls (Adams & Rollings, 2007) because they are familiar to the player and their behavior is known. However, the visual outlook should be different from the corresponding widgets in an ordinary application interface (Adams & Rollings, 2007).

Sometimes it is necessary to show some information outside the gameworld, but in these cases dialogs or user interface widgets should be presented according to the visual theme of the game instead of using the device's visual theme. Any user interface widget which breaks the visual theme, will disturb the player's immersion in the game (Ermi & Mäyrä, 2005; Adams & Rollings, 2007).

Similar heuristics:

- "For PC games, consider hiding the main computer interface during game play." (Federoff, 2002)

GU4: Indicators are visible

The player should see the information such as the current state of the game and the status of the avatar that is required for being able to play the game. Information that is frequently needed should be visible to the player all the time – if possible.

The basic principle of user interface design is to communicate information from the system to a user. In productivity software design, this usually means that all available information is presented in the user interface. In user interface design for a video game, it is not as straightforward as the game mechanics contain lots of information that would be useful for a player to know, but showing all to the player would probably ruin the player experience. Therefore, special attention should be paid to recognizing what kind of information is feasible to show without decreasing the experience.

The key task in game interface design and information presentation is in deciding what information should be shown. This is a specific situation because not all information is equally important all the time (Palm & Koivisto, 2004; Schell, 2008). There are permanent and temporary information needs.

A permanent information need is related to being able to play the game. The status of the avatar and whose turn it is to make the next move (Falstein, 2003) are typically considered permanent information. The player will be frustrated if the avatar dies suddenly and the cause of death is a shortage of energy which could have been recharged with energy packs that the avatar has in his inventory (Falstein, 2003).

A character portrait is normally a small window showing a face of the avatar, a non-player character, or a member of the player's party (Figure 16). The character portrait can be used for building identification between the player and the avatar or party members and to convey more about the personalities of the non-player characters (Adams & Rollings, 2007, p. 253-254).

Figure 16. Character portraits of the player's avatar (Screenshots from Lord of the Rings Online, Turbine, 2007).

A player should also be aware of metagoals, if such goals exist in the game. The metagoals will define how the game goals can be achieved (Malone, 1982). Two basic methods for including metagoals into game design are scoring and time. Scoring as a metagoal means that a player tries to get as high (or low) a score as possible. Correspondingly, time as a metagoal refers to how long it takes to achieve a basic goal. The player can try to achieve the goal as quickly as possible or the player may to try to extend the play session for as long as possible.

Gameplay critical information can be presented to the player using either character strings (text and numbers) or graphical elements (e.g. icons, power bars, brightness regions) (Adams & Rollings, 2007). Graphical elements are commonly preferred because they increase visual appeal, but the most important rule is that the information should be presented clearly enough regardless of the presentation format.

Similar heuristics:

- "A player should always be able to identify their score/status in the game." (Federoff, 2002)
- "A player should always be able to identify their score/status and goal in the game." (Desurvire et al., 2004)
- "Game should provide feedback on the state of the game." (Kiili, 2005)

- "All relevant information should be displayed, such as life points, lives, and ammunition." (Schaffer, 2007)
- "Critical information should stand out." (Schaffer, 2007)
- "Provide users with information on game status." (Pinelle et al., 2008a)
- "Users should be provided with enough information about game (status character, level, health, etc.) but also about other players and online friends in order to play in a cooperative manner as in real life." (Papaloukas et al., 2009)
- "Status score indicators are seamless, obvious, available and do not interfere with game play." (Desurvire & Wiberg, 2009)

GU5: The player understands the terminology

The terminology that is used in the game should be understandable and not misleading or unfamiliar to the players. Technical jargon should be avoided. For instance, terminology that is related to the game concept or features that the game needs from the device should be translated into more understandable language.

Nielsen has stated in the original usability heuristics that the terminology of the user interface should be based on the user's language rather than system-oriented terms or technical jargon (Nielsen, 1993). The same rule applies also to game design, but in addition the terminology used in the interface could also be coherent with the game theme because it would reinforce the game experience.

For instance, in network multi-player games, the players connect to a game server in order to join a play session. The label for the user interface widget is typically "Connect Server." However, from the player's viewpoint, connecting to a server is not the thing, but rather joining the game. A more intuitive label for the widget would be "Join Game" or even "Enter Battle Arena".

Text is a powerful feedback element and it can be used in many different ways in the game interface (Adams & Rollings, 2007). The typical use of text is to provide feedback and explain the meaning of different indicators. Text can also be used for narration, dialog or detailed information about game items (Adams & Rollings, 2007, p. 254).

Similar heuristics:

- "Menu item names should be intuitive and obvious." (Schaffer, 2007)
- "Don't rely on players' memory: Don't use abbreviations or acronyms." (Schaffer, 2007)

GU6: Navigation is consistent, logical, and minimalist

Navigation consists of the game menu and the gameworld. The game menu consists of settings and selections for the desired game session. Different functions should be organized reasonably and possibly on different screens. However, long navigation paths should be avoided. Short navigation paths provide more clarity and are easier to remember. In the main game menu, the player should be able to start a game and have access to other important game features. In the gameworld, navigation should be intuitive and natural. Regardless of the complexity of the gameworld, players should be able to navigate there smoothly. With a proper set of control keys, navigation can be very intuitive and almost invisible.

The game user interface needs to fulfill different kinds of navigation needs. Shell menu navigation follows similar principles and recommendations that are used in menu designs (e.g. Jacko & Salvendy, 1996; Kim, Jacko, & Salvendy, 2011). As the menu contains setting dialogs and links to start a play session, short navigation paths should be preferred as they simplify the menu structure and are easier to remember. Special attention should be paid to how a player can start a play session.

A more challenging task is to design navigation in the gameworld. The gameworld can be either a 3D world, which allows complete freedom of movement, or a simplified presentation in 2D. With a proper set of control keys, the navigation in the gameworld is intuitive for players.

Similar heuristics:

- "Navigation is consistent, logical and minimalist." (Desurvire & Wiberg, 2009)

GU7: Game controllers are consistent and follow standard conventions

Using common conventions in game controllers reduces the time that is needed to learn to play the game since the player can use his or her knowledge from other games. Game controllers usually have specific keys for certain actions and every game should follow them. The game should also provide alternative game controllers because the players should be able to select their preferable controllers.

Games are played on multiple platforms which all have their unique input devices. Depending on the platform, the player can use a keyboard and a mouse, specialized game controllers, or touch screens.

Common input conventions can reduce the time that a user needs to learn a new software application (Nielsen, 1993). This is also true in game design: standard control keys reduce the time that is required for learning to control the avatar and play the game. The player can use his or her acquired knowledge from other games.

Game controllers usually have dedicated keys for certain actions and the game design should comply with these conventions. When using a keyboard as the main input device, moving a game avatar is typically done by pressing arrow keys or 'w', 'a', 's', or 'd' keys to move an avatar forwards, left, backwards, and right respectively. For game consoles, the control key 'x' is used for giving the primary action or selection in the game world, while the control key 'o' is used for a secondary action or cancelling the selection or navigating back.

Birk and Mandryk (2013) have shown that game controllers affect the players' perception of the game and it should be possible to select their preferable game controllers. Some controllers are better suited for certain game genres. Cairns, Li, Wang and Nordin (2014) studied the influence of controllers on immersion. They concluded that naturalness of the controls influences the game experiences especially if there is a natural mapping between the controller and the interaction style. For example, driving a car by using a steering wheel is more immersive than pressing keys on a keyboard.

Similar heuristics:

- "Controls should be customizable and default to industry standard settings." (Federoff, 2002)
- "Mechanics/controller actions have consistently mapped and learnable responses." (Desurvire et al., 2004)
- "If industry standards exist for the controls on the type of game you're working on, adhere to them. For example, if most fighting games use the back button to block, then you should do the same thing." (Schaffer, 2007)
- "Game controls are consistent within the game and follow standard conventions." (Desurvire & Wiberg, 2009)

GU8: Game controls are convenient and flexible

Novice players usually need only a subset of the controls when they start playing the game. On the other hand, veteran players often appreciate shortcuts and more advanced commands. It should be possible to customize the game controls or use shortcuts or macros. However, using shortcuts should not provide a major edge in a competitive player vs. player game. The configurability and amount of controls needed to play the game should be kept at a minimum, but they need to be sufficient. In addition, the controls should be designed according to the device's capabilities.

The customization of game controls is one way of making the learning phase of a new game shorter. The customization can mean simple actions like swapping the left and right mouse buttons which is useful for left and

right handed players. In more complex cases, the player is able to reassign certain actions to particular keys and buttons of the input devices or to access certain menus quickly during gameplay. Another purpose for customization is when a game designer introduces a new control design. The designer should allow the player to customize the controls in case the player does not like it (Adams & Rollings, 2007) or use some earlier control design in case the player does not want to learn new things.

In key reassignment, it is important to make sure that the setting dialog shows the currently assigned actions to the controls, available actions and unassigned controls (Adams & Rollings, 2007). In addition, it would be useful to allow the players to create multiple profiles with saved customizations. This will help the players to manage the game control settings and make them useful.

There are different needs for customization. Allowing the customization of game controls will help players to suit the game to their gaming style (Adams & Rollings, 2007). In the beginning, a subset of controls is needed to start playing a game. Simple game controls help in learning basic actions and getting familiar with the game. Later on, advanced players usually want to customize input devices to suit their gaming style better or there is a need to do complicated tasks, which needs shortcuts or more advanced commands (Adams & Rollings, 2007). However, customization of the input devices or control mappings should not give major advantages over other players in competitive PvP (player vs. player) games (Adams & Rollings, 2007).

Similar heuristics:

- "Minimize control options." (Federoff, 2002)
- "Player should be given controls that are basic enough to learn quickly yet expandable for advanced options." (Desurvire et al., 2004)
- "User interface and controls of the game should be easy to use and learn." (Kiili, 2005)
- "The video game should allow players to customize the settings so that the game accommodates their individual needs." (Papaloukas et al., 2009)
- "Player should be given controls that are basic enough to learn quickly yet expandable for advanced options." (Desurvire & Wiberg, 2009)
- "Player is given controls that are basic enough to learn quickly, yet expandable for advanced options for advanced players."(Desurvire & Wiberg, 2009)

GU9: The game gives feedback on the player's actions

A good user interface has a low response time on the player's actions. An action can be either a single key press or a more complicated input sequence. The player should notice immediately that the game has recognized the action by providing feedback. The most common way of providing feedback is to present it graphically. Other alternatives are to use audio or tactile feedback. Providing only auditory feedback is not acceptable since a player may be playing the game without sounds. Although the game needs to respond immediately to the player's actions, the consequences of the action can be shown to the player later. If an action cannot be performed immediately, the game should notify the player about the delay.

Many games require fast reactions from the player and the player needs to make decisions about the next actions in a fraction of a second. A low response time and immediate feedback is a sign of good user interface design and implementation (Adams & Rollings, 2007). Games are one of the few application domains which use versatile feedback modalities. Graphical feedback is the most common way of presenting feedback, but other modalities such as audio and tactile feedback are used constantly. If audio does not play a major role in the game, a player may mute the game completely.

Although response to the player's actions is usually provided immediately, the consequences of those actions can be represented later. In game design patterns, this is called a *delayed outcome* (Björk & Holopainen, 2005). If the feedback is delayed, the game should notify the player with an appropriate method.

There are two main types of feedback that the player needs to receive. The most important feedback is related to the player's avatar and its current status (health or currently selected item). The feedback should be clear, easily understandable and most importantly, the player should get the needed information with a glance (Adams & Rollings, 2007).

The second type of feedback is related to the player's current location in the gameworld. If the gameworld is not completely visible to a player, there needs to be an indicator for it. This indicator is called a mini-map or a radar screen (Figure 17) and it shows a miniaturized version of the gameworld, or a portion of it (Adams & Rollings, 2007). This will help players to orientate themselves in the gameworld.

Figure 17. A minimap shows the player and other interesting things in the vicinity (A screenshot from Lord of the Rings Online, Turbine, 2007).

Similar heuristics:

- "Game should provide unambiguous and immediate feedback on a player's actions." (Kiili, 2005).

GU10: The player cannot make irreversible errors

The game UI should confirm actions that can cause serious and irreversible damage, which affects the player's ability to play the game. Such errors are typically done in a shell menu or in the setting dialogs to the game character or the player's progress in the game. When mistakes happen, it is helpful to enable recovery.

Making mistakes and errors is part of human nature and they can stem from various reasons. In case an action causes serious or irreversible things in the game, it should be confirmed from the user before executing it. In many cases, computer applications can provide different recovery methods for a user and a possibility to undo harmful actions.

In game user interfaces harmful actions are related to the player's avatar or progress in the game. Sometimes it is possible to delete the avatar, an important game item, or reset the player's progress in the game. In these cases enabling a recovery mechanism is helpful.

Undoing actions in the gameplay is very rare and there is only certain games such as single player puzzle games which allow reverse actions and trying a different choice. The reason for not providing an undo possibility

in gameplay is that it will affect the game balance adversely (Adams & Rollings, 2007).

This heuristic is commonly misinterpreted by the inspectors. It should be noted that the heuristic applies in situations where errors are related to the mistakes that a player makes with the game user interface, not in the gameplay.

Similar heuristics:

- "Provide means for error prevention and recovery through the use of warning messages." (Federoff, 2002)
- "Make it hard to accidentally hit the wrong button. The more trouble hitting the wrong button causes, the farther that button should be from the normal game controls." (Schaffer, 2007)
- "Player error is avoided." (Desurvire & Wiberg, 2009)

GU11: The player does not have to memorize things unnecessarily

The game should not stress the user's memory unnecessarily, unless it is part of the gameplay.

Nielsen (1993) has stated that an application should not stress the user's memory unnecessarily. In game design, the same principle is also applicable and it is not advisable to require a player to remember too many things at once (Adams & Rollings, 2007). Instead, the user interface design should provide a way to look up information that the player needs. If the information is needed constantly, it should be provided as an indicator on the screen (Adams & Rollings, 2007).

The interpretation of the heuristic has sometimes been problematic for inspectors. It should be noted that the heuristic is related to the user interface of the game. It is perfectly possible that one of the game mechanics is memorizing things and the heuristic should not be used to evaluate gameplay.

Similar heuristics:

- "Player's attention should not be distracted by irrelevant things." (Kiili, 2005)
- "Don't rely on players' memory: Players shouldn't have to memorize the level design (but it's arguable there are exceptions)." (Schaffer, 2007)
- "Don't rely on players' memory: Don't require the player to count resources like bullets and Life."(Schaffer, 2007)
- "Don't rely on players' memory: Don't use abbreviations or acronyms." (Schaffer, 2007)

- "Cognitive Workload: The cognitive workload which is not connected to the game play (i.e. connection with the acquisition of skills, the view, the screen orientation and the input methods) should be minimized." (Köffel & Haller, 2008)
- "The game does not put an unnecessary burden on the player." (Desurvire & Wiberg, 2009)
- "Abbreviations should not be used. The players should not be asked to count resources like bullets and life and they should not have to memorize the level's design. Area maps should be easy to learn and should be intuitive to use." (Papaloukas et al., 2009)

GU12: The game contains help

Players do not often read manuals. Instead, the game should teach the player what he or she needs to know to start playing the game. This can be done through a tutorial mode at the beginning of the game. The tutorial mode should be divided into chapters that teach a couple of things at the beginning. Ideally, the tutorial could be embedded completely in the game so that help would be provided every time when it is really needed. Help is also often needed in error situations. If the game provides useful error messages, the player can understand better what caused the problem.

In every game a player needs to learn two things about the game: what the rules of the game are and how to operate game items. There are different methods of how these two things are learnt. The most common way is that a player learns the game by playing the game. A fellow player can also teach and explain how to play or the player can find out operational rules and the meaning of game items from a document or the internet. The document can be provided in an electronic format or it can be a hardcopy manual. Johnson and Wiley (2003) state that the absence of an online manual in a game disturbs the players' immersion in the game as they have to leave the game environment and seek information from a hardcopy manual.

Modern video games have gradually done away with written rules in favor of having the game itself teach players how to play through interactive tutorials (Schell, 2008). The tutorial should be a part of the actual game and it should be entertaining and rewarding. Dividing the tutorial into chapters that teach necessary skills and concepts that the player needs during the first play sessions is preferred. Ideally, the tutorial would be completely embedded in the game so that help is provided every time when it is really needed (Schell, 2008).

The second use of help is in error situations. The game should always provide an error message so that the player understands what caused the problem and how to recover from it.

Similar heuristics:

- "Do not expect the user to read a manual." (Federoff, 2002)
- "Provide an interesting and absorbing tutorial." (Federoff, 2002)
- "Players do not need to use a manual to play game." (Desurvire et al., 2004)
- "There is an interesting and absorbing tutorial that mimics game play." (Desurvire et al., 2004)
- "Players should be given context sensitive help while playing so that they do not get stuck or have to rely on a manual." (Desurvire et al., 2004)
- "Provide instructions, training, and help." (Pinelle et al., 2008a)
- "Player does not need to read the manual or documentation to play." (Desurvire & Wiberg, 2009)
- "Players should be given context sensitive help while playing so that they are not stuck and need to rely on a manual for help." (Desurvire & Wiberg, 2009)

5.2 GAMEPLAY HEURISTICS

Gameplay heuristics include aspects that relate to game content and how the game mechanics are implemented. This is probably the most important module of the heuristic set because these aspects will define what the game will be like.

The heuristics have been published in Paper I.

GP1: The game provides clear goals or supports player-created goals

The players need to understand the goals that exist in the game. The goals can be either set by the game or created by the players. The game should contain both short-term and long-term goals. Short-term goals provide repeated opportunities for reinforcement and keep players motivated to play the game. Long-term goals are usually more difficult to achieve and they can consist of several short term goals.

The most admitted fact among game researchers and game designers is the existence of goals in games. Malone (1982) noted in his early studies that in order for a computer game to be challenging, there must be a goal the achievement of which is uncertain, and the path of achieving the goal should be non-trivial (Adams & Rollings, 2007; Fullerton et al., 2008). The goals of a game are defined by the rules. The game designer can interpret rules anyway she likes (Adams & Rollings, 2007). Achieving the goal means that a player satisfies the designed requirements of the goals and then achieves a particular outcome (Fullerton et al., 2004; Schell, 2008).

The players should be motivated to reach the goal and be able to understand what they are trying to achieve in the game (Clanton, 1998; Falstein, 2003; Barry, 2005). Flow theory notices that a clear goal is the core of an enjoyable experience (Csikszentmihalyi, 1991).

The interesting part of achieving goals is that players can use different tactics and strategies to achieve them (Clanton, 2000). The goals will direct actions towards outcomes (Barry, 2005), and the players can take different actions to achieve the goals. However, it is important that the players understand the purpose of their actions because otherwise they cannot proceed with any certainty (Schell, 2008).

There are different criteria for good goals and not all goals are equally good. Game researchers and designers have presented different qualities for good goals:

- **Skill-based**: players use skills that are taught to them previously in order to achieve a goal (Malone, 1982).
- **Concrete**: Players understand and can clearly state what they are supposed to achieve. The goal should also represent something that the player can easily identify (Malone, 1982; Schell, 2008).
- **Achievable**: The goal does not necessarily have to be achievable (Adams & Rollings, 2007), but players need to think that they have a chance of achieving the goal. If the goal seems impossible to them, they will quickly give up (Schell, 2008).
- **Rewarding**: Players appreciate that the goal is rewarding to them before they try to achieve it. With the right level of challenge, achieving a goal itself is rewarding. To make a goal even more desirable, the player should get something valuable upon reaching the goal (Schell, 2008).

In addition, Malone has presented, based on the work by Morozova, that a goal should not be arbitrary, but it should be a part of the game story (Malone, 1982).

In simple games, it is possible that a game has only one goal, but often a game has multiple goals. In simple games the goals should be obvious to the player (Malone, 1982). In more complex and longer games, there should be both short-term and long-term goals. Multiple goals will make the game multifaceted because players can select which goals they pursue. Achieving different goals may also result in different outcomes in the game.

The balance of the goals will make the players feel that they know what to do immediately and what is the ultimate goal they are trying to achieve (Schell, 2008). Short-term goals support and motivate the players to play the game (Reeve, 2003). Examples of short-term goals can be completing a level or quests or finding a specific item in the gameworld. Long-term

goals are more difficult to complete. They can also include short-term goals. A typical long-term goal is the completion of the story of a game. Another long-term goal could be the development of an avatar to be as good as possible. The distinction between short-term goals and long-term goals in the game should be clear (Schell, 2008). It is possible that some short-term goals are unrelated to the ultimate goal of the game (Barry, 2005), but it is preferable that the players understand how they are related to one another (Schell, 2008).

Goals do not have to be formally structured by the system, but the players can also create goals for themselves (Malone, 1982; Barry, 2005). Some games encourage player-created goals, and some allow the players to choose their goals from a set of pre-defined goals (Björk & Holopainen, 2005). Basically every opportunity for choice or action is an opportunity for goal setting (Barry, 2005). Achieving goals that the players have set for themselves can be very rewarding.

Player-created goals can relate to the meta-goals of a game, the game avatar, or to the gameworld. A player-created goal could be, for example, providing information about a badge that makes a player curious enough to pursue it. This is a player-created goal for a meta-goal of the game.

Similar heuristics:

- "There should be a clear overriding goal of the game presented early. " (Federoff, 2002)
- "There should be multiple goals on each level." (Federoff, 2002)
- "Provide clear goals, present overriding goal early as well as short-term goals throughout play." (Desurvire et al., 2004)
- "Game should provide clear main goal at the beginning." (Kiili, 2005)
- "Game should provide clear sub-goals at an appropriate pace." (Kiili, 2005)
- "It should be clear what's happening in the game. Players should understand and be able to identify goals." (Schaffer, 2007)
- "The game goals are clear. The game provides clear goals, presents overriding goals early as well as short term goals throughout game play." (Desurvire & Wiberg, 2009)
- "New genre games need special equipment and in some cases suggestions are required on how to use it more efficiently. The goal of the game must be clear, so the players do not feel confused." (Papaloukas et al., 2009)

GP2: The player sees the progress in the game and can compare the results

The players should have enough information so that they can see their progress towards the goals in the game. The progress can be shown to the player explicitly or implicitly. The players feel more motivated if they can compare themselves to other players or their previous achievements.

To keep up motivation for playing the game is not only by achieving goals, but it is also the progress which happens on the way (Malone, 1982). Determining when a goal is actually achieved is a major source of satisfaction (Malone, 1982). Most players expect objectives by which their progress can be measured (Barry, 2005). There are two ways to show progress to the player:

- **Explicitly**: The player can see the outcome of their actions in the game with clear numbers or by other means which are visible on the user interface.
- **Implicitly**: There are changes in the gameworld. Depending on the game, the progress can be visible in the visual appearance of the gameworld, the behavior of the non-player characters, or in the availability of game items.

A player's motivation can be increased by enabling them to compare their current performance to previous achievements or against other players. It is also important to give proper feedback, because otherwise the player's performance can feel uninvolving and unimportant (Reeve, 2003). Traditionally, comparisons between players has been done by high-score lists, character levels, rankings of the players, or titles.

Similar heuristics:

- "Game should provide feedback on progression toward goals." (Kiili, 2005)

GP3: The players are rewarded and the rewards are meaningful

The players should receive a meaningful reward as they progress in the game and complete goals. In addition, the reward should be adjusted to the challenge that the player had to face in order to get it. The rewards schedule should be varying and frequent, but still unpredictable.

The players should be rewarded for their progress in the game. Usually the rewards are points, new game items, levels or abilities that will help the players to progress in the game. Another important issue with rewards is that they should be meaningful to the player. An ability to perform an

action 20 percent more efficiently is a less meaningful reward compared to the situation that the game provides a completely new ability, which would allow new tactics and strategies for a player.

The reward should be adjusted to the challenge that the player faces during the gameplay (Adams & Rollings, 2007). There is a possibility of a mismatch and disappointment, if the reward is not what the player expects it to be.

A well-balanced and changing reward schema motivates the players to pursue the next reward even harder instead of giving the reward every time (Williams, Nesbitt, Eidels, & Elliott, 2011). Rewarding the player should be unpredictable but still frequent, because that will maintain the motivation. However, a varying schedule for rewards in the major milestones (for example, completing a level) is not recommended. The rewards can have various forms. The rewards can be something tangible that helps the player to play such as giving more resources or unlocking new levels or areas in the game (Björk & Holopainen, 2005, p. 185-187). On the other hand, rewards can also be something intangible such as a strategic advantage in the game.

Risk and reward are key parts of any kind of competitive gameplay (Adams & Rollings, 2007). "A risk must always be accompanied by a reward. Otherwise the player has no incentive to take the risk" (Adams & Rollings, 2007, p. 27). Uncertainty is a key element of risk. If the player knows the consequences of actions, there is no risk involved (Adams & Rollings, 2007). Uncertainty is often produced by chance, but hiding information is also useful, when it is revealed after the player has taken actions and exposed to a risk. Some games do not have hidden information or the element of chance, but not knowing what other players will do in the game will produce uncertainty (Adams & Rollings, 2007).

Risk and reward mechanism make gameplay more exciting (Adams & Rollings, 2007). Similarly, interactivity is a key element for players in the form of entertainment as the players will complete challenges and practice different skills.

There are different ways to reward the player. Players should appreciate that the goal is rewarding before they have achieved it, so that they are inspired to attempt to achieve it. The player's expectations should not be overinflated because if the player is disappointed with the reward, they will not play again (Schell, 2008). It is also rewarding if the goals can be achieved in more than one way. This will promote rich and dynamic gameplay when players are able to do different kinds of things and achieve goals (Schell, 2008).

Similar heuristics:

- "The game should give rewards." (Federoff, 2002)
- "One reward of playing should be the acquisition of skill." (Federoff, 2002)
- "The game should give rewards that immerse the player more deeply in the game by increasing their capabilities (power-up), and expanding their ability to customize." (Desurvire et al., 2004)
- "Rewards and Feedback: the amount of reward in relation to the effort the player does. Positive feedback, a varying rewarding scheme, valuation of goodness of the player." [translated from Finnish] (Karvonen, 2005)
- "Game should support skill development and provide rewards from development." (Kiili, 2005)
- "The game gives rewards that immerse the player more deeply in the game by increasing their capabilities, capacity or for example, expanding their ability to customize." (Desurvire & Wiberg, 2009)

GP4: The player is in control

The players want to be in control of what is happening in the gameworld. The players should be able to decide on actions they want to take and these actions should have an influence on the gameworld. If full control is not possible, the game should provide at least an illusion of control to the player.

In video games, players usually want to be in an active role and are able to control game events and the game avatar (Adams & Rollings, 2007). Players should be able to decide desirable actions and these actions should influence the gameworld. According to Adams and Rollings (2007) the player can accept some uncontrollable situations which are related to the cinematic sequences of the game, but the game should never seize control of the game avatar and make it do something that the player does not want to do. Nielsen (1994b) has stated that a product (e.g. a game) should not include random uncontrollable events, or tedious or difficult input sequences.

In some games, it is not possible to provide full control of the game events. For example, in gambling games, the player cannot select which numbers win. Instead, the game should provide an illusion of control to the player, even though this is not true. The probability of winning is the same regardless of whether the player selects the numbers or they are randomly selected by the system. However, letting the player select the numbers will make the game much more interesting and entertaining.

Another type of control is related to skills. Players will spend a lot of time in learning new skills in a game. Giving players a chance to enjoy their

sense of mastery of the skill before challenging them with a tough obstacle or opponent is a source of enjoyment (Falstein, 2005). Including variations of the type and difficulty of challenges and actions the player must accomplish is also a part of the feeling of control (Falstein, 2005).

Similar heuristics:

- "The Player has a sense of control over their character and is able to use tactics and strategies." (Desurvire et al., 2004)
- "Create at least an illusion that player is in charge in deciding the progress of the game." (Kiili, 2005)
- "Players should feel in control, so they need the time and information to respond to threats and opportunities. That is, players should see enemies, obstacles, and power-ups coming." (Schaffer, 2007)
- "Level of Automation: The player should be able to execute all actions relevant to the game by herself." (Köffel & Haller, 2008)
- "The players have a sense of control and influence onto the gameworld." (Desurvire & Wiberg, 2009)
- "Players feel in control." (Desurvire & Wiberg, 2009)

GP5: Challenge, strategy, and pace are in balance

The game should be designed so that the challenge is comparable to the players' current skills, then the players do not feel frustrated or bored with the game. In single-player games, the player can often choose the difficulty level and thus affect the challenge. The players learn new strategies as they play the game. There should not be dominating strategies for any part of the game. The pace should be adjusted to the game style and it can be intensive or deliberate. The game should allow the player to take a deep breath once in a while.

Gameplay is defined in terms of challenge and actions (Adams & Rollings, 2007). Challenge is an essential part of game design that captures the players' interest towards the game (Schell, 2008). In addition, Malone (1982) says that goals and challenges are captivating because they engage a person's self-esteem and success in a challenging activity can make people feel better about themselves.

A challenge is any task set for a player that is non-trivial to accomplish (Adams & Rollings, 2007). The most challenges are direct obstacles of achieving a goal, although the game might include optional challenges as well (Adams & Rollings, 2007). A good challenge does not necessarily mean increased difficulty of achieving the goal (Adams & Rollings, 2007).

Overcoming a challenge and performing actions must require certain skills (Adams & Rollings, 2007; Schell, 2008) that a player has, or skills that

can be achieved by playing the game. Most games do not require just one skill from a player, but a blend of different skills. According to Schell (2008, p. 151) the skills required by games can generally be divided into three main categories:

- **Physical skills**: These skills involve strength, dexterity, coordination, and physical endurance. Effectively manipulating a game controller is a kind of a physical skill, but there are also games that require a broader set of physical skills from players.
- **Mental Skill**: These include the skills of memory, observation, and puzzle solving. Almost all games require some mental skills.
- **Social skills**: Typically social skills are related to making friends and influencing people, but the range of social and communication skills in games is much broader. A player needs to be able to read an opponent, fool an opponent, and coordinate actions with teammates.

A game should be designed so that the player's skill level is a good match to the difficulty level of the game. The players should not feel frustrated or bored with the game, but sufficiently challenged (Schell, 2008). According to *Flow Theory* (Csikszentmihalyi, 1991), a game (or an activity) feels frustrating when the difficulty level is higher than the current skill level. Correspondingly, a game (or an activity) might feel boring if the current skill level is higher than what is needed to achieve the goal. Although frustration in relation to skill levels is not recommended, Salen and Zimmerman (2004, p. 348) note that challenge and frustration are essential to game pleasure. Players need to struggle through conflicts and pleasure does not emerge unless the players cannot overcome adversities.

Another aspect related to challenge is an uncertain or varying outcome. There should be some kind of doubt whether a player can win or lose in a game (Malone, 1982). There is no challenge or reason for competitive play, if the end result is obvious from the beginning (Malone, 1982). Malone pointed out that an uncertain outcome can be achieved in four different ways: adjusting the difficulty level, randomness of events, information hiding, and having multiple levels of goals (Malone, 1982). The right balance between outcomes and skills depends on the game style.

In single-player games, the player can often select a desirable difficulty level for the next play session. Schell (2008) notes that chance means uncertainty and surprises in games, which is an essential part of fun. Chance should give positive feelings of excitement and challenge rather than feelings of hopelessness and lack of control (Schell, 2008). The player should always feel a fairness of the outcome. Information hiding is a complicated task because the game should provide clear performance feedback to enhance the challenge and not reduce self-esteem too much to discourage a player (Malone, 1982). An uncertain and varying outcome of

goals means that some goals can be evident while their influence on the long-term goal might be uncertain.

The players learn and apply new strategies as they play a game (Rollings & Adams, 2003) and rich and dynamic gameplay is achieved when players are able to do different kinds of things and achieve goals (Schell, 2008). Strategic options do not need to be obvious in the game, but the designer needs to be aware of how a player expects the game to operate (Barry, 2005).

The most critical aspect in the players' strategies is that there should never be an option which is significantly better than the others, because the players will always pursue that option. It also reduces the players' willingness to explore alternative strategies in the game (Rollings & Adams, 2003; Fullerton et al., 2008; Schell, 2008). This is called a dominant strategy. Rollings and Adams (2003, p. 244) define that "a dominant strategy is one that surpasses all others by being the best one to choose under any circumstances".

Clanton (2000) says that a game should provide progression of challenges and a sense of pacing. If the game requires that the player thinks about the next moves in the game, there should be time for doing that. In intensive gameplay, time to analyze different options should be kept minimal because otherwise it will slow down the gameplay. In any case, the game should set the pace so that the challenge level increases sometimes faster and sometimes slower (Falstein, 2005). The slower challenge level increase allows the player to take a deep breath once in a while during the play sessions. Egenfeldt-Nielsen et al. (2008, p. 121) state that different games have certain conventions of pace and speed that is required from the players to be successful in the game.

Similar heuristics:

- "Pace the game to apply pressure to, but not frustrate the player." (Federoff, 2002)
- "There must not be any single optimal winning strategy." (Federoff, 2002)
- "Game play should be balanced so that there is no definite way to win." (Federoff, 2002)
- "Play should be fair." (Federoff, 2002)
- "Design for multiple paths through the game." (Federoff, 2002)
- "Pace the game to apply pressure but not frustrate the player. Vary the difficulty level so that the player has greater challenge as they develop mastery. Easy to learn, hard to master." (Desurvire et al., 2004)
- "Game play should be balanced with multiple ways to win." (Desurvire et al., 2004)

- "Player experiences fairness of outcomes." (Desurvire et al., 2004)
- "Challenge should match a player's skill level." (Kiili, 2005)
- "It should be clear what's happening in the game. Players should understand and be able to identify failure conditions (How they lose)." (Schaffer, 2007)
- "The game is paced to apply pressure without frustrating the players. The difficulty level varies so the players experience greater challenges as they develop mastery." (Desurvire & Wiberg, 2009)
- "Challenge, strategy and pace are in balance." (Desurvire & Wiberg, 2009)
- "The game is balanced with multiple ways to win." (Desurvire & Wiberg, 2009)

GP6: The first-time experience is encouraging

The first impression of the game is formed within a few minutes and it is very difficult to change. The players should feel that they have learned the basics and have accomplished something. The first play session should make the player desire the next play session.

The first impression of a game is formed within the first five minutes of playing. Later on it is very difficult to change it. Along with the first impression, the players could assume that they have a reasonable chance of succeeding, they have accomplished something and they are rewarded (Rouse, 2001; Mulligan & Patrovsky, 2003). The first play session should make the player desire the next play session. A negative first-time experience can discourage players and make them ignore the game in the future. The risk in game development is that the game is designed to be too hard from the beginning (Rouse, 2001). Recently, Cheung, Zimmermann and Nagappan (2014) have studied first hour experiences and state that the game design should not only provide enjoyment during the first moments of the game, but show how the players will enjoy the game in the future. This can be achieved by providing anticipated elements, things that a player encounters later in the game, or show the depth of narrative and gameplay with interesting game mechanics.

The first play session is also a learning phase of the game. A player should learn the basic skills that are needed to play the game. The game should introduce the first achievable goals and involve the player in the story, if such an element exists in the game. If learning new things is hard and the player has difficulties in playing the game at the beginning, the first-time experience may be frustrating.

Similar heuristics:

- "Provide a gentle on-ramp: the game should be easily learned." (Clanton, 1998)
- "Get the player involved quickly and easily." (Federoff, 2002)
- "The first player action is painfully obvious and should result in immediate positive feedback." (Desurvire et al., 2004)
- "Get the player involved quickly and easily with tutorials and/or progressive or adjustable difficulty levels." (Desurvire et al., 2004)
- "Upon initially turning the game on the player has enough information to get started to play." (Desurvire et al., 2004)
- "The first ten minutes of play and player actions are painfully obvious and should result in immediate and positive feedback for all types of players." (Desurvire & Wiberg, 2009)
- "Upon turning on the game, the player has enough information to begin play." (Desurvire & Wiberg, 2009)

GP7: The game story, if any, supports the gameplay and is meaningful

Even though the story plays an important role in many games, it should not dominate the gameplay. Some games do not even have or need a game story. If the game has a story, it should fit the other elements in the game and sound plausible to the player. Dialogue with non-player characters (NPC) should be meaningful and interesting to the player.

Fantasy is a strong source of enjoyment (Malone, 1982), and many games have integrated story elements or storytelling in the game (Adams & Rollings, 2007). If the game has a story, it usually has a huge impact on goals, the gameworld, and the avatars. The story will bind them together and make things meaningful in the game (Schell, 2008). The basic principle is that the story should sound plausible and fit the other elements in the game. The story also directs players and gives them a context within which their experiences fit (Barry, 2005). The story should make a player feel that he is inside the story and can affect the flow of events (Adams & Rollings, 2007). When this happens it is called narrative immersion, or imaginative immersion (Ermi & Mäyrä, 2005), where the player is completely involved and accepts the gameworld and the events of the story (Adams & Rollings, 2007).

There are two basic methods of how stories are integrated into games. The most dominant method in video games is commonly called the "*string of pearls*" or cinematic scenes. The idea is that a completely non-interactive story is presented in the form of text, a slideshow, or an animated sequence (Schell, 2008). Between these cut-scenes, the player controls an avatar and performs tasks to achieve goals in the game. When some specific goal is achieved, the game shows the next cut scene which reveals

how the story continues. This method enables a balance between interaction and storytelling in the game (Schell, 2008).

Even if the method is criticized for taking away control from the player (Falstein, 2005), it allows the player to enjoy a finely crafted story which is seamlessly integrated into the game (Schell, 2008). The critical element is the length of the cinematic scenes. It is preferable to use shorter scenes which do not take control away from the player for too long (Falstein, 2005).

Another method is a story machine in which a game generates a sequence of events that is interesting and relates to something else (Schell, 2008). The story is created as the player plays the game.

Avatars are often very closely related to the game story. Interesting avatars reflect their traits in their interaction with other characters and main characters preferably have some interpersonal relationships (Schell, 2008). In addition, good avatars are also meaningful to the player and there should be a reason to care about them (Fullerton et al., 2008). Falstein (2005) notes that the control of the avatar is also an important aspect and the avatar should not act autonomously and take control away from the player. Avatars should be designed so that they look alive or change their behavior according to the events in the game. Players might find avatars repulsive, if they try to mimic humans, but fail in some way (Schell, 2008).

There are a couple of risks related to the story elements in games. The story should never dominate the gameplay, but remain in the background. Another risk is the multiple choices the story enables for the players. A problem arises when some of these choices become meaningless and end up in the same outcome or conclusion (Schell, 2008). Instead, the players should make their own decisions in the game and the story should follow the players' choices. If the story has multiple choices, the player may want to see the different options and play some sequences again, which can be frustrating and cause confusion about whether he has been playing the wrong track all along. This will diminish the player experience.

Finally, it should be noted that a game does not even have or need a story to be fun (Fullerton et al., 2008). Gameplay itself can create a story of victory and loss. Sometimes it can be useful to let the players create their own stories and the game just provides a background story, on which the players can base their own stories (Mulligan & Patrovsky, 2003).

Similar heuristics:

- "Create a great storyline." (Federoff, 2002)
- "Player understands the story line as a single consistent vision." (Desurvire et al., 2004)

- "Players discover the story as part of game play." (Desurvire et al., 2004)
- "Provide consistency between the game elements and the overarching setting and story to suspend disbelief." (Desurvire et al., 2004)
- "Player is interested in the story line. The story experience relates to their real life and grabs their interest." (Desurvire et al., 2004)
- "Player is interested in the characters because (1) they are like me; (2) they are interesting to me, (3) the characters develop as action occurs." (Desurvire et al., 2004)
- "Storyline: Originality, Plausibility, Surprise, Perfection, The story is interwoven into the game, the Storyline follows the game." [Translated from Finnish] (Karvonen, 2005)
- "Use frame story to situate challenges to meaningful context." (Kiili, 2005).
- "Use frame story to integrate challenges to clear entity and to support perception of goals." (Kiili, 2005)
- "Game story encourages immersion (If game has a story component)." (Desurvire & Wiberg, 2009)
- "There is an emotional connection between the player and the gameworld as well as with their avatar." (Desurvire & Wiberg, 2009)

GP8: There are no repetitive or boring tasks

A game should not require repetition of tasks without changing any conditions. Often, this repetition happens when the player needs to reach a certain goal before the game becomes interesting or challenging. However, during the training phase (tutorials), it is useful to repeat certain tasks so that the player learns and practices, for example, how the character is controlled in the game.

Executing tasks is a fundamental part of any game. The flow of tasks is guided by goals and players progress in the game by completing tasks. The number of tasks is usually enormous and the biggest pitfall in game design is that the tasks start to repeat after a while.

Adams and Rollings (2007) say that tasks (or challenges as an alternative term) can be unique, recurring or continuous, but the game design should change some conditions of how the tasks can be completed. If the tasks are repetitive, it will make the gameplay boring (Fullerton et al., 2008).

Completing repetitive tasks in the game is also known as grinding or tread milling (Sorens, 2007). To recognize grinding, players will repeatedly use the same set of actions and strategy over and over again to complete a certain goal. Often grinding happens before the game offers new and interesting content or more challenging tasks for the player.

However, it should be said that completing tutorials is not grinding even though there might be some similarities. During the training phase, the player needs to repeat certain actions or tasks multiple times, for example, to learn how to control the avatar in the game.

Similar heuristics:

- "Include a lot of interactive props for the player to interact with." (Federoff, 2002)
- "Player's fatigue is minimized by varying activities and pacing during game play." (Desurvire et al., 2004)
- "The players finds the game fun, with no repetitive or boring tasks." (Desurvire & Wiberg, 2009)
- "Any fatigue or boredom was minimized by varying activities and pacing during the game play." (Desurvire & Wiberg, 2009)

GP9: The players can express themselves

The players should be able express themselves by, for instance, customizing their characters, acting in a certain way, or modifying the gameworld. Allowing the players to customize and personalize their game characters makes it more probable that they feel attachment to a game.

Players are creative creatures and a game often inspires them to express themselves in the gameworld. This can happen by customizing the game avatar, acting in a specific manner in the gameworld, or even modifying the gameworld. The game design should be balanced for the need of the players to behave within the system with their desire for expression (Barry, 2005).

Customization or personalization makes it more probable that the players feel togetherness with the game avatars and are attached to the game (Björk & Holopainen, 2005; Fullerton et al., 2008). The simplest action is to allow the player to name the game avatar. The players will also like to create or modify the character or other game items (Adams & Rollings, 2007). If it is not possible to modify the character, the character design becomes even more important (Björk & Holopainen, 2005). Modifying the gameworld increases the feeling of ownership to the player. Changing textures of the game items or the game logic requires time and artistic skills, but it is also very rewarding. These kinds of activities are often called "modding".

Expressive play can also affect the gameplay directly or it can be purely cosmetic (Adams & Rollings, 2007). However, the interesting aspect in expressive play is that players do not always make rational choices even if the choice can affect the gameplay. Players may select game items that

they like regardless of the consequences. This is an indication of how strong the appeal to self-expression is (Adams & Rollings, 2007).

Similar heuristics:

- "Allow players to build content." (Federoff, 2002)

GP10: The game supports different playing styles

The players vary a lot in terms of experience and preferred playing style. One of the main differentiating factors is the players' attitude towards risk-taking. The story can also influence whether the player wants to play as a hero or a villain. The player types will also determine how the players prefer to interact with the gameworld and with other players. In very simple games, different play styles are usually not supported and all players will have only one role.

Players differ quite a lot in terms of preferred playing style and experience. Further, their motivation to play a game can influence preferred playing style (Yee, 2006). Players will always confront challenges in the game, but their attitude towards risk-taking varies (Adams & Rollings, 2007). Some players have a tendency to a more aggressive and inherently risky approach while others may prefer a more defensive approach in which they try to minimize the risk. A game should be designed so that both styles are possible and one style does not have an advantage over the other.

The story may affect the preferred playing style. The player should have the possibility to select the initial role of the avatar to better suit the player's playing style (Fabricatore et al., 2002). Sometimes the player may want to play the heroes of the game, but similarly they could prefer an opposite role for their avatar such as villains. However, it should be remembered that the player or his avatar is a protagonist of the game regardless of whether he plays a hero's or a villain's role (Clanton, 2000) and the gameplay should be adapted accordingly.

Some games can support more specific playing styles. Bartle's (1997) player types are commonly used for categorizing players in massively multi-player online role-playing games (MMORPG). The players' preferences to interact with other players and with the gameworld define their player type. Basically, the interaction style is either collaborative ("act with") or dominative ("act upon"). Bartle (1997) has introduced four player types:

- Socializers, who prefer to socialize with other players.
- Achievers, who like to compete with the game mechanics.
- Explorers, who wish to explore different aspects of the game.
- Killers, who enjoy dominating other players.

Although Bartle's player types are defined for MMORPG games, they are applicable to other games, as well. Most games can be evaluated based on their ability to support exploring and achieving. Socializing and killing activities are commonly included in all kinds of multi-player games. Depending on the game mechanics, multiplayer games can direct or limit players towards certain playing styles. For example, in cooperative games players tend to play different characters to complement each other and achieve synergies of the characters (El-Nasr et al., 2010).

Exploration is a particularly interesting aspect in games because it has the power to provide other types of enjoyment (Adams & Rollings, 2007). Games can permit actions that are not required to surmount a challenge or the player can have extended interaction with game items. The player can also play with the avatar or non-player characters and try unusual movements. Playing with avatars which have ragdoll physics (Witkin & Baraff, 1997) is one example of the common practice of exploring how realistically game items behave. Fullerton et al. (2008) list additional categories for playing styles which describe the pleasures of play from the player's point of view. According to them, players can also act as collectors, craftsmen, storytellers, directors, or competitors, to name a few.

Many games enable the player to explore the gameworld beyond the paths that are needed to complete the game objectives (Adams & Rollings, 2007). Especially if the gameworld resembles the real world, the players are curious to see how far away it is possible to travel. These are interesting factors to increase player enjoyment and player experience.

The need to support different playing styles mainly concerns more complex games where the gameworld and the story give possibilities for different styles. In very simple games, different play styles are usually not supported and all players will have only one role.

Similar heuristics:

- "The game supports a variety of game styles." (Desurvire & Wiberg, 2009)

GP11: The game does not stagnate

The players should always feel that it is possible to reach the goals and the game progresses towards the goals. Game items or the balance of power should not counterbalance each other to result in an infinite loop of meaningless actions. The game should recognize immediately when the game is over and inform the players. The ending of the play session should be clearly indicated and restarting the game should be possible.

Fullerton et al. (2008) mention stagnation as one of the problems in design which can deteriorate the player experience. The players should always feel that the game progresses towards the goals and it is possible to reach the goals. Sometimes the players might not recognize the progression because it happens in too small steps or the required actions towards the goal are not clear (Fullerton et al., 2008).

Another type of stagnation is the balance of power or game items. A certain combination of game items can cause a situation where it is impossible to win the game because the game items have equal opposite strengths and they counterbalance each other.

Finally, the end state of the game should be recognized instantaneously and it should be indicated to the players that the play session has ended. The player should be offered a possibility to restart the play session.

Similar heuristics:

- "Avoid lengthy dead ends. It is unfair to need to know the end of the game to play the beginning correctly." (Clanton, 1998).
- "Don't make it easy for players to get stuck or lost. The goal of the game and the next step towards that goal should always be clear. There should be a sense of progress towards that goal, so players never feel lost or like they're going around in circles." (Schaffer, 2007)

GP12: The game is consistent

The gameworld and actions should be consistent and logical to the player. If something works in the beginning, the player assumes that it also works later on. Correspondingly, if the player is able to perform a certain action in the gameworld or for a game item, the player assumes that a similar kind of action is possible for other similar objects or in a similar situation as well. More actions can become available as the player progresses through the game.

The players operate in the gameworld and based on that they make assumptions about what is possible and what is not. If the gameworld looks like the real world, the player might assume that real world principles apply in the gameworld and actions can be completed in a similar way. The gameworld should always be consistent (Rollings & Adams, 2003). Consistency in the gameworld means that certain actions can be done throughout the play sessions. Adams and Rollings (2007) note that actions that are normally available to the player do not change from level to level. The only exception is that more actions can become available as the player progresses through the game (Adams & Rollings, 2007, p. 341).

Rules are definitions and instructions that players agree to accept for the duration of the game. Rules establish the goal of the game and the meaning of the different activities and events that take place within the magic circle (Adams & Rollings, 2007). The term magic circle means that the gameworld is restricted spatially, socially and culturally from the rest of the world (Huizinga, 1955).

Similar heuristics:

- "Provide consistent responses to the user's actions." (Pinelle et al., 2008a)

GP13: The game uses orthogonal unit differentiation

Each game item should have a purpose in the gameworld and it should be notably different from other similar game items and preferably restricted in some way. If the player needs to select character classes or roles in the game, they should be functionally different. In addition, interaction between the avatar and the non-player characters should support different interaction styles.

Usually the game design provides lots of game items and avatars for a player to choose from. Smith (2003) has introduced the term "Orthogonal Unit Differentiation" to describe possible problems related to game items. The term means that game items and other units in the game are designed to be functionally different or they are used for different purposes. If a player has to choose between three game items (e.g. weapons), but they work equally well to accomplish a goal, the choice is functionally meaningless (Falstein, 2005). When game items are functionally different, it encourages strategic play and expands the set of possible actions to a player.

The same strategy applies to an avatar. If the game provides a possibility to select a different type of avatar to play, the player should understand the role of the avatar and the avatar's abilities should be coherent with its role (Fabricatore et al., 2002). Otherwise, the choice of the avatar is meaningless.

The choice of game items can also be influenced by their availability. Providing a limited supply of game items, we can create tension, challenge and provide an opportunity for different strategy choices in the game (Barry, 2005). Game items should not be useless, readily or infinitely available (Barry, 2005).

Orthogonal unit differentiation also relates to the interaction between the player and non-player characters (NPCs), and what kind of impact the player's actions has. The player might want to interact with non-player characters using different interaction styles (e.g. insult, flatter or ignore). If

the character's response is identical in each case, the choice of the interaction style is meaningless from the player's perspective (Falstein, 2005).

Similar heuristics: None

GP14: The player does not lose any hard-won possessions

The game should maintain the possessions that the player has earned while playing the game and the player cannot lose them accidentally. However, in some cases the game can provide very high risks and the player can stake valuable game items which can be lost during the gameplay.

Usually players have to work hard to get some specific game items, develop their avatar or achieved certain things in the game. The simplest form of hard-won possession is awarded points and ranking on the high score list. The game design or interaction design should ensure that the player does not lose the avatar because of a single mistake in the game or a random event that kills the avatar after several weeks of hard work and dedication (Fullerton et al., 2008).

The game design should not discourage the players from attempting something in the game (Falstein, 2003) because of the risk of losing something valuable in the game. Sometimes the game design may break this rule on purpose and provide a more exciting gameplay experience by offering a very high risk. However, in these cases, the player should have a possibility to win back any game items that have been lost previously (Falstein, 2003).

In some games, the player can purchase game items with real money. In these situations, game design should be implemented so that it is not possible lose these items accidently because of poor user interface design or because of the game rules.

The design rule was originally presented by Falstein (2003) in the Game Developer Magazine. Later on the rule has been renamed to "*Don't Penalize the Player*" in Falstein's 400 project rule list (Falstein & Barrywood, 2006). This could be a better name for the heuristic, but for the sake of clarity and consistency with other publications I have kept the original name.

Similar heuristics:

- "If there are tasks which you expect to be challenging, don't require players to complete them more than once. That is, make sure that if they die soon after completing a hard task that they don't have to complete the hard task again." (Schaffer, 2007)

- "The players should not lose any hard won possessions." (Desurvire & Wiberg, 2009)

5.3 MULTI-PLAYER HEURISTICS

The playability heuristics presented in this section are focused on the social activity and the community-building mechanisms that are common in multi-player games. There are two kinds of heuristics. The first eight heuristics are independent and consider multi-player aspects. The ninth heuristic "MP9: Players should play with comparable players" is a supplement to gameplay heuristic GP6 because the first time experience of players is very much influenced by with whom a player is playing the game.

The heuristics have been published in Paper II.

MP1: The game supports communication

Communication is one of the cornerstones in multi-player games and, depending on the game, communication channels for different purposes need to be supported. In-game communication can be either synchronous or asynchronous. Chatting is the most frequently used method for in-game communication, but it can be restricted to cover certain areas in the gameworld or certain players. Asynchronous messaging is used with players who are not currently online.

Communication is one of the cornerstones of multi-player games (Friedl, 2002) and depending on the game, different communication channels need to be supported. Relationships between players are maintained in communities and communities do not exist without communication (Koivisto, 2003).

There are two kinds of verbal in-game communication in a game: synchronous and asynchronous (Koivisto, 2003). Chatting or textual messages are used frequently for in-game communication because it can be used for asking support or advice, social exchanges and coordination, and scheduling activities (Seay, Jerome, Lee, & Kraut, 2004). However, studies have indicated that sometimes players do not notice all messages to fellow players because they are mixed with status messages in the same window (Cornett, 2004).

In-game communication can be visible to everybody or it can be restricted to cover only certain areas in the gameworld or limited to certain players. The players should control with whom they like to communicate and ignore other players, if necessary. Limiting communication can be used to reduce spam or keep the amount of messages reasonable. Furthermore, it can be used to create interesting game mechanics (Davidsson, Peitz, &

Björk, 2004; Björk & Holopainen, 2005). Hiding tactical information can create a need for collaboration between players.

Asynchronous messages are a part of in-game communication and they are often used for delivering messages to players who are not online. They can be either mail messages or message boards which are shown to players when they enter the game (Seay et al., 2004).

Using out-of-game communication, referring to web sites, discussion forums and other messaging tools, are commonly used by power gamers to improve community knowledge of the gameworld, characters, game items and other things which are relevant to players (Taylor, 2006).

Voice communication is less frequently used and it usually needs 3rd party applications (Seay et al., 2004). Mobile devices are designed to foster communication between people. Verbal communication would be an ideal solution for player-to-player communication. It is also faster and more convenient than typing messages. Voice communication was trialed in the mobile game '*Pathway to Glory*[5]'. The players were able to send short voice-messages to other players while playing the game.

Similar heuristics:

- "Appropriate communication tools: provide communication features that accommodate the demands of game play." (Pinelle et al., 2009)
- "A game in a social network should support all the tasks, which facilitate the communication and socializing of players. The game should have "shared" versions or "shared" applications in order to direct "social networking friends" to tasks that enhance socializing." (Papaloukas et al., 2009)

MP2: There are reasons to communicate

Communication is an essential part of social interaction. When players are either collaborating or competing this will generate discussions. Moreover, if players see how other players are doing in the game it will also generate discussion topics. The communication topics usually range from game related topics to common conversation to kill time during boring play periods.

A game should provide meaningful and interesting events which create discussion topics for the players. Typically this means that the players need to either collaborate or compete against each other (Zagal, Nussbaum, & Rosas, 2000). In addition, when a player sees the

[5] http://www.redlynx.com/games/legacy/pathway-to-glory

progression of other players in the game, it will also generate discussion topics.

Topics for discussions can be very different. Players are very likely to discuss strategies, tactics and secret information that are discovered in the game (Koivisto, 2003). Games quite often provide interesting game objects, puzzles, events, difficult boss monsters that require collaboration and teaming up with fellow players and give the players reasons to communicate. Taylor (2006) has noticed in her studies that players often recount fights to other players. Players also use in-game communication to entertain themselves during boring play periods (Taylor, 2006).

Similar heuristics: None

MP3: The game supports groups and communities

The game should support both short-term groups and long-lasting communities because they will increase social interaction and keep players playing the game. Short-term groups need to support communication and care taking of other members of the group. Long-lasting communities are more persistent groups which help players to advance in the game and increase social interaction.

A multi-player game will benefit from features which encourage community formation (Järvinen et al., 2002). Belonging to a community is a motivational factor for players to keep playing the game (Koivisto, 2003; Ducheneaut, Yee, Nickell, & Moore, 2006). The game should enable short-term group formation as well as long-lasting communities.

Temporary groups are often formed to complete tasks and achieve goals that would be impossible to do alone. Groups usually have an in-game communication channel, which helps in coordinating actions and communicating with group members. In addition, group members can usually see some status information (e.g. health and location) of other members, which helps in taking care of teammates (Koivisto, 2003). The group members may or may not know each other beforehand, but there are usually some rules which everybody should follow. A successful group may continue to play for several guests (Nardi & Harris, 2006).

Longer-lasting communities are usually formed by a group of players who want to create a more permanent alliance or a guild. The guild does not have to have a specific goal other than a possibility for the players to have a group to identify with (Nardi & Harris, 2006). However, it is a common practice that the guild organizes activities for its members (Koivisto, 2003) and helps them to advance in the game (Ducheneaut, Yee, Nickell, & Moore, 2007). The guild usually has hierarchical ranks and roles for guild members (Koivisto, 2003; Nardi & Harris, 2006) that are needed to run the guild. Sometimes there are also persons who do special tasks as

playmakers in the community (Salovaara, Johnson, Toiskallio, Tiitta, & Turpeinen, 2005). In addition, guild members often help each other in ways that do not happen in temporary groups (Koivisto, 2003).

Similar heuristics:

- "Community formation should be supported." (Kiili, 2005)
- "Collaboration and Communication: The interpersonal communication and collaboration should be supported by the entirety of the game (such as game play and setup)." (Köffel & Haller, 2008)
- "Support coordination: provide features that allow players to coordinate their actions during cooperative game play." (Pinelle et al., 2009)
-

MP4: The game helps the player to find other players and game instances

Teaming up with other players is an essential aspect in multi-player games. The players should have a sense of the presence of other players and be able to find them. The game can direct players with similar levels to the same area in the gameworld where the players will meet each other. Alternatively, the game can automate matchmaking and help players find other players or game instances.

There are different possibilities for interaction between the players ranging from player vs. player to multilateral competition (Fullerton et al., 2008), but first a player should have a sense that there are other players around and be able to find them and game instances. Especially new players often feel unsecure and inexperienced with the game and are willing to join a group of other players (Cornett, 2004). Making new contacts and friends is an important activity in an online gameworld because the player does not necessarily know other players beforehand (Taylor, 2006).

In online multi–player games, the players should be able to see if their friends are online and find them in the gameworld. The game could also inform the players when their friends get online.

If players do not know others in advance, the game could provide means for meeting up and getting to know new players in the gameworld (Koivisto, 2003). The game could also provide a search feature, which allows the player to use character properties or titles to search other players. Some games feature an automatic group-forming process, which makes things easier (Cornett, 2004).

Sometimes the player needs to find a proper game instance to join the game. It is preferable if the player is directed to meet players with a

similar level range (Koivisto, 2003). Allowing players to join the game as spectators before actually joining the game can help in finding appropriate game instances.

Similar heuristics:

- "Flexible matchmaking: provide matchmaking features to help people find players with similar interests." (Pinelle et al., 2009)
- "Simple session management: provide session management support that allows players to start new games, and that allows them to find and join appropriate games." (Pinelle et al., 2009)

MP5: The game provides information about other players

Knowing other players' online status, level and rank will help in planning and coordinating actions in the game. The information can be used when deciding whether to interact with some player and if the player can be trusted. Being aware of other players in the gameworld will increase the social aspects of the game.

Knowing basic information such as the level or rank and the online status of another player is useful information in order to plan and coordinate actions in the game and find others to group and play with (Ducheneaut et al., 2006). The information can also be used when deciding whether to interact with another player and if a player can be trusted (Jensen, Davis, & Farnham, 2002).

Being aware of other players in the gameworld will also act as a social aspect in the game. Other players provide an audience and a social presence for the player (Ducheneaut et al., 2006). The audience makes success or loss in the game more meaningful to the players. Multi-player games are also reputation games and the players like to show others what they have achieved in the game and showcase their latest accomplishments (Ducheneaut et al., 2006; Chen & Duh, 2007).

Social presence can be achieved by providing in-game communication methods which overcome normal boundaries. For example, the guild chat channel can be used for conversation between the guild members regardless of their location in the gameworld (Ducheneaut et al., 2006).

Similar heuristics:

- "Meaningful awareness information: provide meaningful information about players, including information about action, location, and status." (Pinelle et al., 2009)

MP6: The design overcomes the lack of players and enables soloing

There are always situations when other players are not available. There may not be enough players in the gameworld or a player might play the game at times when there are not many other players around. There might also be some personal reasons to play alone. The player might have insufficient language skills to communicate with other players or they might want to try to survive by themselves in the game, especially on the lower levels. Whatever the reason is, a multiplayer game should also provide content for solo players.

Although multi-player games are intended to be played with other players, there are situations when this is not possible or even wanted. Many online games have a "critical-mass problem", which means that there are not always enough players in the gameworld to make the gameplay meaningful. Many players may also want to play multi-player games alone. Especially at lower levels players prefer soloing because it provides a faster progression through levels and most of the game content is soloable on these levels (Ducheneaut et al., 2006).

There can be personal reasons for this behavior. Some players are shy or have insufficient language skills for an international community. The players might also play the game at times when there are not many other players around. Group formation and waiting other players to show up can take some time (Ducheneaut & Moore, 2004).

The playing style might also restrict the availability of fellow players. Sometimes it is not easy to find players who are willing to collaborate on tasks that require teamwork or a specific way of playing the game (e.g. strict role-playing and character building). Sometimes players may want to test the game mechanics and see if it is possible to complete specific tasks or levels without help from other players.

Whatever the reason is, the game design should take these into account and provide content for solo players even if it is a multi-player gameworld (Pardo, 2006).

Similar heuristics:

- "Reduce game-based delays: minimize interaction delays by reducing temporal dependencies between players." (Pinelle et al., 2009)

MP7: The design minimizes deviant behavior

The game design should minimize deviant behavior of the players, which is a common problem in multi-player games. Some players try to purposefully disrupt other players' gaming experience with their own behavior. Players can cheat,

exploit or hack the game system or they harass other players intentionally. Game designers should pay special attention to minimizing deviant behavior and implement mechanisms that prevent such actions.

It is unfortunate that antisocial behavior is a common problem in multi-player games. Some players purposefully disrupt other players' gaming experience with their own behavior. A player's behavior is deviant when it is not in accordance with the community standards (Bruckman, Curtis, Figallo, & Laurel, 1994). Deviant behavior may occur in multi-player games as cheating, exploiting, hacking, and grief play.

Cheating can be defined as "an action by a player that violates the rules of the game 'as written' or commonly understood" (Chen, Huang, Huang, & Lei, 2006). Exploiting is usually the result of not finalized game mechanics which allow the players to do unintended actions and it can be considered as a cheating technique among players. Hacking is an intentional action to create an exploit or look for loopholes in the implementation. Intentional harassment of other players is called griefing and players utilize the game structures and physics in unintended ways to cause distress to other players (Warner & Raiter, 2005). Grief play is usually targeted towards less advanced players or players whose avatars are in vulnerable states (Foo & Koivisto, 2004).

Foo and Koivisto (2004) have listed different ways of griefing: harassment, power imposition, scamming and greed play. Harassment means that a griefer's intention is to cause emotional distress to the victim and it often happens through communication. Power imposition means that a player kills other players without direct benefit, it accompanies verbal abuse, or the act is repeated several times. Scamming is often related to transactions between the players in exchange for game items. Greed play implies the player's motive to benefit his actions regardless of the annoyance caused to others.

The game designers should pay special attention to minimizing deviant behavior in a game and implement mechanisms that prevent such actions (Baughman & Levine, 2001). Preventing grief play often requires restricting player-to-player interaction. However, careful considerations should be taken because too strict restrictions may lead to dull gameplay. The probability of bad behavior is smaller if the players know each other well or the game design encourages positive interaction (Jensen et al., 2002).

Similar heuristics:

- "Manage bad behavior: provide technical and social solutions for managing cheating and unsavory behavior." (Pinelle et al., 2009)

MP8: The design hides the effects of the network

In online games, the effect of the network may become an issue for positive player experience. Updates to the gameworld should always happen without delays, but latency can disrupt the gameplay and cause jitter in real-time interaction. As the players can move in the physical environment while playing mobile games, they can be unintentionally disconnected from the game. The game design should minimize disturbances to the player or other players in the game.

Multi-player games can be played either in close proximity or at remote locations. In games where the players are remote, network latency or disconnections may become an issue.

Latency can disrupt the gameplay and cause delays, especially to real-time interaction. Latency critical aspects in the game are movement and special actions such as combat behavior (Dick, Wellnitz, & Wolf, 2005; Fritsch, Ritter, & Schiller, 2005). Typically, the players can experience that other players' avatars are jumping around or freezing in one location and it is difficult to coordinate future actions because it is not known where other players will be in a moment. There are different techniques developed to overcome latency problems (Gautier & Diot, 1998; Pantel & Wolf, 2002; Li, Li, & Lau, 2004; Chen et al., 2006). Mobile games can be very sensitive to latency problems, but the problem can be hidden (Palm & Koivisto, 2004; Xu, 2008).

As the players sometimes move in the physical environment while playing mobile games, it is possible that they are disconnected from the game unintentionally. Therefore, the game system should handle these disconnections gracefully and minimize the disturbances to the player or other players (Nokia, 2003).

Similar heuristics: None

MP9: Players should play with comparable players (Supplements GP6)

In multi-player games the player population can be diverse. Some players have played longer than others and are thus more experienced in the gameworld and the tactics. They also have more developed avatars. The game design should help novice players to get familiar with the game and let them practice their skills and develop their avatar apart from the more experienced players during the first play sessions.

In online multi-player games, the players' experience and skill levels can be very different and some players might have played longer than others. This usually gives them tactical advantage in the game. The first play session of the novice players can be spoiled by more experienced players, especially in games in which player vs. player (PvP) conflict is a core game mechanic.

All players should feel fairness in the beginning of the game (Adams & Rollings, 2007). Therefore, it is advisable to separate novice players from more advanced players and let the novice players get familiar with the gameworld, practice their skills, and compete with fellow players who are approximately on the same skill level.

Similar heuristics: None

5.4 MOBILITY HEURISTIC

This module contains heuristics which are relevant to games that are played on mobile devices in varying contexts. There are two kinds of heuristics. The first three heuristics are independent heuristics that consider mobile games. The next four heuristics are supplements to the game usability heuristics because playing games on mobile devices will modify some game usability aspects.

The heuristics have been published in Paper I.

MO1: The play sessions can be started quickly

Starting a play session should be quick and easy because it is often used to fill up micro-breaks. The player should be able to skip introduction sections or startup screens. The game menu should be designed so that frequently used actions are not hidden behind long navigation paths. The default settings for the control keys and UI customization should be feasible for most of the players. In addition, any changes to the settings should be saved. The player should be able to continue play sessions from the point where the player stopped during the previous play session.

Starting a play session is many times a spontaneous activity which is used to fill up a micro-break (spare time between two planned activities or tasks) (Anttila & Jung, 2006; Cui, Chipchase, & Jung, 2007). Play sessions in mobile games are typically shorter than in computer or console games. Micro-breaks can last from 30 seconds to a few minutes (Cui et al., 2007). Continuous attention to the mobile device is fragmented and broken down to bursts of interaction, which can last only seconds (Oulasvirta, Tamminen, Roto, & Kuorelahti, 2005). Therefore, play session should be started quickly and let the players use the available time to play the game.

Entering the gameworld quickly can be achieved by a careful user interface and system design. Introduction sections or other startup screens are very common in computer and console games. These are used for loading and transferring the player into the gameworld or advertising the game developer or other game titles. In mobile games, long intros are not recommended since they take too much time. There can be an introduction but the player should have a possibility to skip it (Nokia, 2006).

The shell menu is another critical step for entering the gameworld. Games usually contain multiple settings for customizing the game user interface or giving information about the player or input devices. If these default settings are good for most players, the player will be able to access the game faster. In addition, the game should save any changes that the player makes to the settings (Nokia, 2006).

Finally, the time of entering the game can be shortened by enabling the player to resume the previous game state and continue playing immediately.

Similar heuristics: None

MO2: The game accommodates the surroundings

Playing a game in mobile contexts should not disturb non-players in the vicinity. A game should provide means for conveniently adjusting the volume level or muting the game. Alternatively, the game should respect the profile settings of the device or ask whether the player wants to disable the audio features of the game.

Computer and console games are usually played indoors where a disturbance generated by the game is minimal. Mobile devices, on the other hand, are played in varying contexts (Bertini et al., 2009) and possibilities for generating disturbances for non-players are multifold.

Noise is the most common way of disturbing other persons in the vicinity. A game should provide a means for conveniently adjusting the volume level or mute the game completely. If the volume level controls or settings are not easily accessible, the game could ask at the beginning if the player wants to disable the audio features of the game (Nokia, 2004a). Finally, the game should always respect the profile settings. If the device is, for instance, in silent mode, the game should be muted automatically.

From the interaction point of view, the game should never use audio as the only method of providing feedback to a player unless the game concept is based on audio. Typically the game might be muted or the player is not able to hear the sounds in the current usage situation (Studio 7.5, 2005). The use of headphones cannot be expected.

Similar heuristics: None

MO3: Interruptions are handled reasonably

There can be both internal and external interruptions, which will disturb the gameplay. Internal interruptions include incoming calls and received messages. External interruptions are related to the player's task switching in the current context or other unexpected events, which cannot be anticipated. If an interruption happens, the game should pause automatically and allow the player to continue gaming later on.

Some game researchers say that a possibility to pause a game and thus interrupt event-time decreases the aesthetic experience of playing the game (Juul, 2005). Since mobile devices are designed to be multi-purpose devices and they are used in different contexts, interruptions while playing a game are inevitable. The context will create a great challenge for providing a smooth gaming experience and there are different contextual factors that affect the gaming experience (Korhonen et al., 2010).

There are multiple reasons both external and internal why a play session can be interrupted. The external interruptions stem from the communication capabilities of the device. The player might receive a call or somebody could send a message. Incoming calls are usually handled immediately or at least the device switches to call mode and shows who is calling. For incoming messages, there can be a small notification and the message can often wait until the play session is over. The player's current context may also require the player's immediate attention and the player needs to quit the play session and do something else.

Internal interruptions are related to the technical capabilities or the limits of the mobile device. It is possible that while playing the game, the player moves unintentionally out of network coverage and is disconnected from the game server without a warning. The device might also run out of battery in the middle of the play session and the game ends suddenly.

To ensure a positive gaming experience, mobile game design should ensure that interruptions and continuation of the play session is as smooth as possible. Single-player games should implement a pause functionality, which enables pausing the game at any time and continue the play session later (Nokia, 2005a) or the game is paused automatically when interruptions happen. The game state should be saved frequently and the player should be able to resume at the same point (Dixon, Mitchell, & Harker, 2004).

In multi-player games, this is not that straightforward, as there are other players involved. However, interruptions will happen in multiplayer games as well and the player may need to stop playing the game for a short period of time. In these situations, the game could replace the interrupted player and control the avatar with artificial intelligence (Nokia,

2003). Another option is that the player is removed from the game and the game continues with the remaining players.

Similar heuristics: None

MO4: The graphical design is accommodated to current brightness (Supplements GU1a)

The game settings should provide means to compensate changes in the environmental context. The players should be able to adjust screen brightness and color scheme in the game if necessary.

Mobile games are played both indoors and outdoors where the environmental context (Korhonen et al., 2010) can change unpredictably. A game should be playable in changeable lighting conditions. Bright sunlight directly to the screen can make seeing things on the screen difficult. The game should recognize the environmental context and adjust screen brightness or the color scheme accordingly to maintain legibility (Studio 7.5, 2005). It is advisable to use high contrast with colors rather than similar hues.

Similar heuristics: None

MO5: The player should be aware of some device features while playing (Supplements GU3 and GU4)

The player should be aware of certain device functions when playing games with a mobile device. The most important information is network connection, battery, and modality of the keypad. Showing a clock will help the player to estimate when it is time to end the play session. Information should be presented using the same graphical style and user interface widgets that are used in the game user interface.

Although the device user interface and the game user interface should be separated, there are some device functions that should be visible to the player in the game interface. The most important information is network connection, battery, modality of the keypad and a clock.

There might be a sudden loss of connectivity that will affect gaming and therefore, the player should be able to quickly check the connectivity status. Games will typically drain battery due to constant network connectivity or the use of sensors (e.g. location sensor, GPS) and a player should be aware of how much battery is left. In case of a mobile device with a physical keypad, a keypad indicator is needed to show whether it is in number, alphabetical or specific text entry mode (e.g. T9 mode). The clock is not as critical as other status information, but it is mainly a convenience feature and knowing the time would help the player to know when it is time to end the play session.

If the game shows status information notifications related to these indicators, the same graphical style and user interface widgets should be used to present the information to the player.

Similar heuristics: None

MO6: Mobile devices have their own conventions for input (Supplements GU7)

Game design should follow the input conventions of the target device. For mobile devices with a physical keypad the keys have specific functions like a selection key. For touch screen devices, the conventions are related to gestures.

Mobile devices may not be as flexible as most of the other game devices and the possibilities to customize game controls are often limited (Nokia, 2006). Mobile devices have existing input conventions which should be followed with mobile game designs as well.

For mobile devices with a physical keypad, some functions are specifically important and assigned to certain keys and they should be accessible even though the device is used for playing games (Nokia, 2006). Mobile devices are primarily used for communication between persons. Therefore, incoming calls during a game session should be managed effectively. There are two keys assigned to call handling and they should never be used for controlling a game (Nokia, 2004b, c):

- The Send key (usually identified with a green symbol) is for answering calls. This will also move the game to the background because in-call functions are activated.

- The End key (usually identified with a red symbol) is for rejecting incoming calls.

In addition, there might be some other keys such as volume or camera keys which are defined dynamically. Following the conventions of using such keys is not always straightforward. On the mobile device's keypad, number five is the selection key. Forum Nokia mobile game usability guidelines specify control keys for mobile games with a physical keypad (Nokia, 2004b, c).

For touch screen devices, the number of physical keys is reduced to a minimum. Usually there is a universal Home key or Back key, which will quit the current application and take the user back to the main menu of the device. Most of the interactions happen through softkeys or with gestures. Typical interaction includes tapping, swiping, or tilting. There are certain conventions in gestures. For example, zooming in and out is usually done with a pinch gesture (Moscovich, 2007). The interaction design for the game should use the same gestures.

One interesting thing about mobile devices is the design driver for operation. If the device is meant to be operated one-handedly, the game should be playable with one hand, as well. This will present a great challenge to designers, because the same hand is used for the interaction and to hold the device (Studio 7.5, 2005). Correspondingly, the device's standard input methods should be used for controlling the game.

Similar heuristics:

- "If possible, users should be able to play mobile games with one hand." (Schaffer, 2007)

MO7: The tutorial should respond to immediate demand (Supplements GU12)

For mobile games, the tutorial should be tailored to present only information that a player needs during the current play session. The tutorial could be divided into several subsections which will span across several play sessions.

The game teaches the player what he or she needs to know to start playing the game. Players do not often read manuals, and a mobile game does not usually even have a paper manual (Jones & Marsden, 2006, p. 233). At the beginning of the game, a tutorial mode is usually helpful for a player to learn to play the game. However, offering a complete tutorial is not recommended because mobile game players usually do not need all of the information when they start playing a game. The first play sessions can be short and the players do not necessarily encounter all aspects that are introduced in the tutorial. Therefore, the tutorial should be presented as the player progresses in the game.

Another aspect is that game design should focus on entertaining the player even though she or he had a couple of minutes to play the game. The player should accomplish something during the first play session.

Similar heuristics: None

5.5 Context-Aware Heuristics

Pervasive mobile games are an expansion of mobile games which will benefit from the mobility aspect of a gaming device. Montola et al. (2009) have presented a comprehensive list of pervasive game genres to study these aspects. In this dissertation, I use the term pervasive mobile game to describe games that use different sensor and radio technologies of mobile devices to collect and share data. At minimum, mobile devices can be used to collect data about the current context and transfer it to the game system (Paavilainen, Korhonen, Saarenpää, & Holopainen, 2009). The location and

availability of WLAN hot spots, GSM base stations and Bluetooth devices are typical examples of data that are used in the games (Sotamaa, 2002; Broll & Benford, 2005; Peitz et al., 2007).

Although we have done only preliminary studies in this field (Paavilainen et al., 2009; Saarenpää, Korhonen, & Paavilainen, 2009), some aspects emerged immediately and they should be seen as factors affecting the playability of pervasive mobile games. The following is an initial list of heuristics (they have been published in Paper III) that could be used to evaluate context-aware games. Developing heuristics for pervasive games has been studied very little and there are only a few studies that have touched upon this topic (e.g. Röcker & Haar, 2006; Jegers, 2008).

CA1: Perception of the current context

Discovering the correct context and playing the game when the context is favorable for a player is one source of fun in pervasive mobile games. The challenge, however, is that the game system and the player may not have a mutual understanding of the current context. In addition, the player may be uncertain of how the game system interprets the current context even though it was clearly observable. The game system should only use information that is collected from the player's current location and avoid rigid thresholds. The game system can also notify the player of what the current context is, if it does not contaminate the game experience.

Using context information and defining conditions in the gameworld based on the context is a key element in many pervasive games. Determining the right context and playing the game when the context is favorable for a player is one source of fun in these games. The challenge is that the player and the game system may not have a mutual understanding about the current context. Context information can be inaccurate or unavailable due to sensor failures, noise and user-supplied information is subject to human error problems (Henricksen & Indulska, 2004). The inaccuracy of information can be due to different measurement points. For example, current weather or temperature can fluctuate within a close range. A player may also be uncertain of how the game system interprets current context information even though it would be clearly observable (Lonthoff & Ortner, 2007).

Therefore, game design should not use rigid thresholds with context information or rely on context information from a single source. The game system should preferably use context information that is collected from the player's current location and possibly using the same device that the player is using for playing the game. This would minimize the possibility of error in different measurement points.

Another option is that the game system notifies the player about the current context. In this solution, the difficulty is in presenting the context information without contaminating the game experience. Most games rely on hidden information that could be essential for a player to succeed in the game and context information belongs to this category in pervasive mobile games. In addition, the current context may be a summary of multiple data that should be informed simultaneously.

Similar heuristics:

- "Offer information about the player's context to the player." (Heikkinen, Strömberg, Koivisto, & Suomela, 2006)

CA2: Players should have an equal chance to play

The utilization of context information in the game system should be carefully designed and all players should have an equal possibility to access relevant information. The dynamic nature of the context or player-related reasons might cause some context information to become unreachable. The game should be designed so that player progression is not dependent on context information that is unreachable by many players.

Including context information into game mechanics should be carefully considered. Game designers need to ensure that all players can somehow access the relevant information in play sessions. The players need to have an equal possibility to play the game even though the lack of time or unfavorable location may cause troubles playing the game successfully and getting the necessary game resources (Bell et al., 2006; Suomela & Koivisto, 2006; Peitz et al., 2007).

As context information is mostly dynamic and uncontrollable, it may place players in unequal positions. Furthermore, there might be player-related reasons why a specific context cannot be reached. The player might be unable to play the game at a specific time of the day. The player's current social context may not permit gaming or there may be other activities ongoing at the same time. It is also possible that some resources are not available at the player's current location and the player is prohibited from traveling to other locations. Montola et al. (2009, p. 107) call this contextual adaptability, which means that the game should adapt to changes in the social environment to minimize negative effects of the gameplay.

The game should be designed so that player progression is not dependent on the context information that may be unreachable for many players. If such design is not possible, another option is to enable the required context information through fellow players. This may also enhance and deepen the social interaction between the players.

Finally, due to the dynamic nature of context and other limitations, the game could also provide alternative tasks for players to complete, if a certain context is not available.

Similar heuristics:

- "Consider how much effort (physical, cognitive, emotional) the players have to make." (Heikkinen et al., 2006)

CA3: Adjustable play sessions

Play sessions of pervasive mobile games are often blended into other activities of the player. This will cause situations where the play sessions are fragmented and the player might have to quit playing the game without prior notice. The game should be designed so that the chance of winning the game is not dependent on the player's ability to play the game. In addition, the player might need to adjust the pace of the game and match it to the available playing time. Sometimes it is preferable to play the game almost in real time while some other times a slower pace is preferred.

Pervasive mobile games will break the continuous play session concept in which playing the game is the main activity and the game is played for a certain period. With pervasive mobile games, it is more common that the players play the game when time permits it. A play session often blends into other activities, which makes gaming a secondary activity. As a result, play sessions can be short and the player might have to quit playing the game without prior notice.

Even though short play sessions will promote the social adaptability of pervasive mobile games (Peitz et al., 2007), they will also set new challenges for game design. It should be possible to leave the game at any time without dramatically reducing the player's chance of winning the game. However, leaving the game might have some temporal influence on the player's abilities or available game items (Montola et al., 2009). In multi-player games, the game design should consider what kinds of effects the players' traffic in and out has on other players.

In pervasive mobile games, the players might also want to adjust the game pace and match it to the available playing time and attention span. Sometimes the player might prefer to play the game in real time while some other time a slower pace is preferred. This will give a possibility to lie back and do other things. The game can also vary how quickly the player needs to react to game events (Montola et al., 2009). The player can keep track of game events, although the game requires only minimal attention. These things are possible in asynchronous gameplay.

Similar heuristics:

- "Support spontaneous and occasional playing." (Heikkinen et al., 2006)

CA4: Communication outside the gameworld (Supplements MP1)

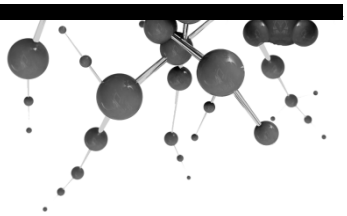
Pervasive mobile games expand the need for communication outside the gameworld. The players need to be aware of game events even though they are offline. Offline communication should not be overwhelming, because it can lead to negligence and disturb the player's other activities. The game events that require the players' actions should be delivered immediately, while others can be postponed and delivered as a summary type of messages.

Pervasive mobile games expand the communication needs outside the gameworld. The players need to be aware of game events even though they are not actually playing the game. The players might also want to communicate with players who are not online while they are playing the game. Linner, Kirsch, Radusch and Steglich (2005) have introduced a framework which enables rapid development of pervasive games that support offline communication. The system buffers received messages and the player can see them on the next login. Lindley (2005) describes a pervasive mobile game that allows the players to receive alerts, although they are not actually playing the game (the game character is in a dormant mode). Players have quite often created their own tools or utilized existing tools, if the game does not support offline communication (Koivisto, 2003).

In offline communication, it is important to decide the frequency and the content of the messages. Normally, the player will be informed about various game events while they are playing the game, but with offline communication it is not feasible to deliver the same amount of information. If the number of messages is overwhelming, it will easily lead to negligence and may disturb the player's other activities. The game design should adapt to the player's playing style. Devoted players should be able to access the messages rapidly and the communication frequency could be higher whereas casual players want to see only the most important messages while they are offline (Montola et al., 2009).

It should also be considered how much delay is acceptable in delivering messages. Some messages should be delivered immediately allowing fast reaction to events, while others can be postponed and delivered more as "For Your Information" type of messages. The player should be able to define the proper number of messages depending on the player's needs and interest in monitoring the game offline (Crabtree et al., 2007).

Similar heuristics: None



6 Studies

The development of a new heuristic set usually requires two phases: heuristic development and heuristic validation (Ling & Salvendy, 2005). In the heuristic development phase, new heuristics are defined by modifying existing heuristics or defining a new set without referencing any existing heuristics. Heuristics are defined by evaluating existing products in the domain using either user testing or other methods to find out usability problems, which are then abstracted to general principles (Dykstra, 1993). In the validation phase, the newly developed heuristics are compared against other heuristic sets or user testing results (Ling & Salvendy, 2005). It is also common that heuristics undergo several modification rounds until the final version emerges (Ling & Salvendy, 2005).

The major part of the dissertation work has involved a series of game evaluations first to define playability heuristics for different kinds of mobile games, and then validating the playability heuristics by comparing them against other commonly used playability heuristics. Some games were evaluated using both the expert review and the playtesting methods to see the differences between the results. Along with heuristic development, the objective of each experiment was to study the expert review method and identify changes or additions that the procedure might require when evaluating games.

Figure 18 illustrates the research process and displays the temporal relationships between the phases. In the development phase, an initial set of playability heuristics was defined and used in mobile game evaluations. The heuristics were refined based on the results of the first evaluations and new heuristics were added to make the set more complete. Experiments consisted of multiple game evaluations. The objective was to

get data about the playability heuristic set in realistic evaluation situations which will then help in developing the heuristics further.

In the validation phase, game evaluations were conducted using the defined playability heuristics set along with another playability heuristic set which was developed by other researchers. In addition, the evaluations were conducted by external inspectors to see whether the heuristics were understandable and usable. The objective of the experiments was to gather data about the strengths and weaknesses of each playability heuristics set, which would then help in developing our playability heuristics set further.

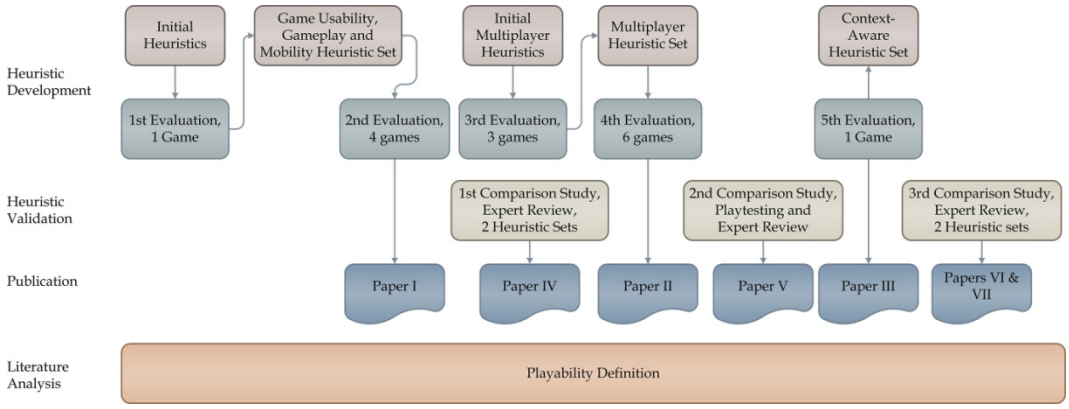


Figure 18. Research process for developing playability heuristics.

In the validation phase, the focus was also on the effectiveness of expert review in comparison to the playtesting method. The objective was to study how the evaluation results differ and whether the expert review method is able to provide data that is comparable to playtesting results and useful for game designers.

6.1 Study I: Playability Heuristics for Mobile Games

The aim of the first study was to introduce the expert review method for game evaluations and describe the development of playability heuristics that will guide the evaluation task of the inspectors. It was clear from the beginning that instead of having a huge list of heuristics, we need to group them somehow. There will be more heuristics than what the general usability heuristic list contains. Categorization also enables easier management of the heuristics during the evaluation because, for certain aspects of the game, only a subset of the heuristics is needed.

We presented a model of playability heuristics and introduced three core modules: *Game Usability*, *Gameplay* and *Mobility* that are common for any (mobile) video game. The Game Usability module contains heuristics that cover the user interface and the controls that a player uses for interacting with the gameworld. The Gameplay module contains heuristics which help in evaluating the game mechanics, story, challenge, goals and other

aspects which are typically present in games. The Mobility module concerns issues that affect how effortlessly a player can enter the gameworld and how the game behaves in diverse and unexpected environments. In addition, mobile games set some new challenges for usability aspects which were described in another publication (Nokia, 2006).

During the first game evaluation, it became apparent that our initial playability heuristic list (11 heuristics) was unable to describe several playability problems in a game. We explored these problems and discovered that our heuristics are lacking important aspects especially in gameplay. There were also playability problems related to the user interface and usability in general for which we did not have a proper heuristic. As a result, we added 18 heuristics into our original heuristic set and ended up with 29 playability heuristics.

In the second evaluation round of the study, we evaluated four games by using the expert review method and the enhanced playability heuristics. The results of the evaluations showed that the playability heuristics were now able to describe the identified problems accurately and none of the problems were left unassigned, except for problems in multi-player features. We also saw the first signs of the usefulness of the method in game evaluations as the heuristics helped the inspectors to focus on specific aspects in the game. Especially gameplay would have been difficult to evaluate without the playability heuristics. The evaluation of the user interface and the mobility aspects of the game resembled a normal usability evaluation of productivity software and they were therefore more straightforward to evaluate.

The detailed description of the study has been presented in Paper I.

6.2 Study II: Playability Heuristics for Mobile Multi-player Games

During the first study, we evaluated three multi-player games and there were several playability problems which concerned the social interaction between the players. In multi-player games, social interaction is an important aspect of the gaming experience and it should be covered in playability evaluations. After the first study, it became apparent that multi-player games need their own set of heuristics.

In the second study, we introduced a new module that would cover social aspects such as presence, communication and interaction in mobile and other types of games. For the development of multi-player heuristics, we followed a similar procedure as in the previous study and first conducted a literature review and defined an initial list of multi-player heuristics. The

list contained six heuristics to cover possible playability problems in the social interaction of the players and the heuristics were added as a multi-player module (Figure 19).

In the first evaluation round, we evaluated three multi-player mobile games that were under development and the evaluation focused especially on the player-to-player interaction. The results indicated that all games contained playability problems related to the multi-player heuristics that we had defined. However, there were also problems that were not covered by the initial heuristic set. We further defined additional heuristics to cover the missing aspects and included them in the module. In the second evaluation round, we explored six commercial PC games and conducted an informal playability evaluation of them to see how well the heuristics would help in identifying multi-player aspects in games that can be found on the market. The refined module contains eight multi-player heuristics and it extends the usefulness of the model as it is now capable of evaluating both multi-player and single-player games.

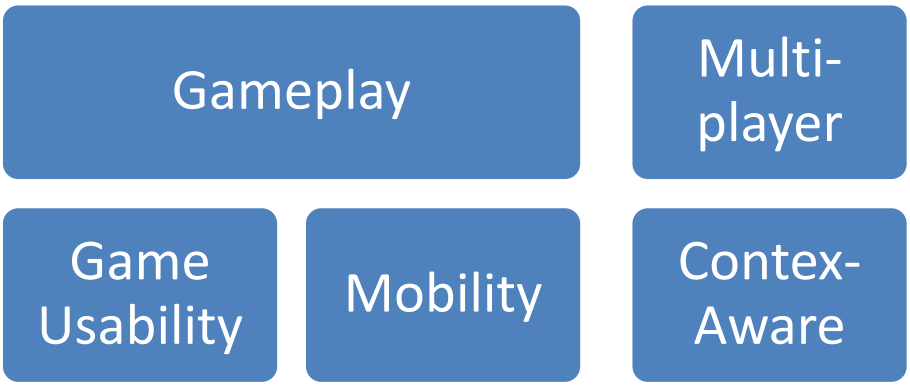


Figure 19 Playability Heuristic Modules.

The study showed that the final multi-player heuristic set addressed many social interaction issues in the games and they can be used to identify playability problems which might make a game less enjoyable. From the methodological point of view, the most interesting finding was that there needs to be a critical mass of players (or inspectors) in order to complete the evaluation successfully. The inspectors cannot conduct the evaluation independently, but they need each other in the game and to play together in order to evaluate the social interaction.

The detailed description of the study has been presented in Paper II.

6.3 Study III: Playability Heuristics for Context-Aware Mobile Pervasive Games

The first two studies developed playability heuristics for games that reside on stationary gaming platforms such as the PC and game consoles. Most

mobile games fall into this category when the gameworld is fully virtual. However, there are mobile games that are not just scaled down versions of the PC and game console games, but implement mobility in a broader sense. Mobile games should utilize some elements which are typical of how and where the mobile devices are used (Järvinen, 2002b).

For this purpose, we implemented an experimental pervasive multi-player mobile game, which would allow us to explore the contextual factors and how they affect the playability of a game. During the study, players played a game for one week and after that we interviewed them about their playing experiences. We were interested in their opinion about the pervasive mobile game in general and their attitude towards embedding contextual information into the game design.

The user study of the game revealed that the context influences the playability of the game and the players' gaming experience quite a lot. Based on the study, we defined four new heuristics that specifically tackle contextual issues in pervasive mobile games and introduced a new module called *Context-Aware* to the model (Figure 19). As context information is used to define some conditions in the gameworld and the players try to discover the correct context to complete their tasks, it becomes important for the players to know the current context and their perception should match the context in the game system. The utilization of context information also requires that all players can access relevant information equally. Pervasive mobile games alter the players' playing behavior as play sessions are mixed with non-gaming activities, but a player would like to keep track of game events. Therefore, communication outside the gameworld becomes an important feature and the game design should carefully consider the frequency and the number of messages that are sent to offline players.

The detailed description of the study has been presented in Paper III.

6.4 Study IV: Comparison of Playability Heuristic Sets with Expert Inspectors

The process of developing new heuristic sets contains a validation step (Ling & Salvendy, 2005). In the validation step, the newly developed heuristics are compared either to traditional usability heuristics or other domain-specific heuristics. In our studies, we decided to use only domain-specific heuristics as a baseline, because, as it was previously noted, traditional usability heuristics are not effective in game evaluations. In the first validation study, our objective was to compare two playability heuristic sets with expert inspectors and to explore the strengths and the possible weaknesses of the heuristic sets.

In the experiment, eight inspectors from the academia and the game industry conducted an evaluation using the expert review method and domain-specific heuristics. The inspectors worked in pairs and two groups used one playability heuristic set. The inspectors played the mobile game for one hour and reported encountered playability problems in the game.

Based on the results, expert review produces valuable results and the heuristics helped the inspectors to conduct the evaluation. However, both playability heuristic sets should be developed further before the game industry can widely adopt them. The baseline set suffered from several problems. There were overlapping heuristics and some heuristics caused difficulties for inspectors to use them effectively during the evaluation, and there were too many of them. In addition, the categorization of the heuristics was confusing for some heuristics. The difficulties of using the heuristics were visible in the evaluation results. The inspectors could identify playability problems quite easily, but assigning a heuristic was harder, and 30% of the problems were left without a clarifying heuristic.

Our playability heuristic set received a more positive welcome from the inspectors. The heuristics were more consistent in wording and abstraction level, and their categorization was better. The main development point is to combine a heuristic with a short description of its meaning because in the study setup the descriptions were provided in a separate document.

From the methodological point of view, the study gave rise to several interesting observations. The inspectors commented that the evaluation task should not be considered a normal play session because the inspectors are focused on the evaluation task and it prevents immersion in the game. Consequently, this means an inability to evaluate the player experience aspects of the game. This observation is consistent with the intended usage of the method because experience studies require different methods. Another observation was that mobile devices can have different technical capabilities which will influence the reported playability problems. Some reported problems were specific to the device that was used during the evaluation. Technical differences specifically influenced the audio features and the visible content on the screen. The third observation was that there are similar challenges in consistently identifying playability problems in games that have been reported in productivity software evaluations (e.g. Molich & Dumas, 2008). This is an interesting observation because the gameworld design is normally quite linear in many games and the players will go through the first levels in a specific order. Therefore, all inspectors should encounter the same aspects in the gameworld and possible playability problems during the evaluation. The evaluator effect might be one possible explanation for this issue because the inspectors had different backgrounds and their experience in

game design and evaluation tasks differed. Other explanations for the inconsistency might be the two different heuristics sets that were used and the insufficient instructions that were given for writing the problem reports. Some pairs mainly reported general problems whereas other teams focused on very specific problems. The fourth observation relates to the origin of the reported problem. It is important to know whether the problem is located in the gameplay or in the user interface. Evaluations of productivity software do not have this type of challenge because the evaluation only covers the user interface.

The detailed description of the study has been presented in Paper IV.

6.5 Study V: Comparison of Playtesting and Expert Review in a Mobile Game Evaluation

Basically, there are two primary evaluation method types available that are used to evaluate the usability of products: usability inspection methods and user testing. In video game evaluations, conducting a user test with players has been the dominating method and usability inspection methods have only been used infrequently. The main reason among game designers is that they want to gain useful feedback from players to improve the overall experience of the game (Fullerton et al., 2008). In many cases, game designers prefer informal playtesting such as self-testing or playtesting with confidants because showing an incomplete game design to outsiders is considered difficult (Fullerton et al., 2008). The expert review method can provide a valuable external opinion to game designers and showing an incomplete design is not a problem because playability experts are used to evaluating different kinds of prototypes.

In this comparison study, expert review and playtesting were used to evaluate a mobile game. The objective of the study was to explore if the inspected playability problems are consistent with the observations from playtesting. The positive results would indicate the effectiveness and usefulness of the expert review method.

Two playability experts evaluated the mobile game. Six players were recruited for user testing. The procedure of the evaluation was similar in both cases. The experts were instructed to explore the gameplay and the user interface. In addition, their objective was to progress in the game as far as they could during the evaluation session. There were no specific scenarios or tasks given to the inspectors. The playtesting was conducted in a usability laboratory and followed the standard procedure. A moderator observed the players' performance in the game and possible playability problems. In addition, another person observed the game session in an observation room and made observation notes. The game

session lasted for 60 minutes in both cases. The expert review evaluation preceded the playtesting.

In total, 46 playability problems were reported in six playtesting sessions. Expert review reported 32 playability problems. The consolidated list of playability problems contained 53 problems and they were related to both the gameplay and the user interface. 25 out of 53 problems (47%) were found by both methods. 7 problems (13%) were solely reported by the playability experts. 40% of the playability problems were playtesting specific. The problems reported by a single method provide an interesting starting point for comparing differences in the evaluation methods.

Both expert review and playtesting focused on the user interface, and 70% of the reported problems were categorized as user interface problems. The players were also more sensitive to encountering problems in the user interface and the playtesting reported 17 exclusive problems. When we look at these problems in detail, in most cases only a single player encountered a problem. Moreover, the problems received either a medium or a minor ranking in severity. This means that the players find some issues disturbing initially, but once they learn the game and get used to the disturbing issue, the annoyance disappears. Examples of minor problems are related to terminology and visualization.

In the gameplay, both players and experts reported exclusive problems. The experts reported seven problems and four problems were reported by the players. During the evaluation, the experts completed more levels than the players and some playability problems were reported on levels that the players never reached. In addition, only experts reported playability problems related to goals and rewards in the game. This can be considered as an advantage of using inspectors in a game evaluation, because finding these problems requires that gameplay is explored analytically and some levels were played repeatedly.

One playability problem that the players reported was a particularly interesting one, because the problem conflicted with the players' experiences with the previous versions of the game. The game designers had made one fundamental change in the game mechanics, which dramatically modified the behavior of the game avatar in the gameworld. These kinds of problems are difficult for the expert review method to capture because they are related to the players' previous knowledge and experiences. A half of the players had difficulties with the new behavioral model of the game avatar, which indicates that it might become a common playability problem.

The results of the comparison study showed that expert review is an efficient method for evaluating playability. Experts usually have the required domain expertise and they might even belong to the target

audience. Their knowledge of different games and genres also helps to identify possibly problematic issues. Another advantage is that the experts can analyze their own behavior and identify problems that are caused by the game design. In playtesting, the players are interested in playing the game and they get easily focused and immersed in the gameplay. In this mental state, they do not diagnose the encountered problems automatically, but instead they try to overcome them and continue playing the game. Even though it is the moderator's responsibility to identify the problems that the players have in the interaction with the game mechanics or the user interface, they need to express them somehow. Quite often the players were silent for ages and played intensively. This observation is consistent with the experiences that game designers have had (Schell, 2008). The third advantage is that the experts are usually skillful players and they will use their time efficiently and the game is evaluated thoroughly. For players, difficult challenges are obstacles that they try to solve and time is spent inefficiently. The players do not try to do other things than try to proceed in the game unless the moderator instructs them to do otherwise. The fourth advantage is that the experts pay attention to issues that might be ignored by the players. During the evaluation, the experts are in an evaluation mode and can explore things like short-term goals and rewards while the players are mainly playing the game. The players do not have an interest in such things by default. The moderator needs to ask the players' opinion about them.

Nonetheless, expert review is not superior to playtesting. Instead, both methods should be considered complementary to each other. Expert review can be used to discover obvious playability problems before conducting evaluations with players and playtesting can reveal problems which possibly originate from the players' previous experience, skills and knowledge. In this study, the experts and playtesting identified playability problems consistently. In addition, expert review was able to identify problems that were not touched in playtesting at all.

The detailed description of the study has been presented in Paper V.

6.6 Study VI: Comparison of Playability Heuristics with Novice Inspectors

In the third validation study, we corrected the experiment design issues and harmonized the mobile devices used for the evaluation and thus got more accurate data on how heuristics are used in the evaluation task. In this study, our objective was to study the use of the playability heuristics when novice inspectors evaluate a game and to explore in detail the effectiveness of the inspector groups and inter-evaluator reliability.

We recruited 36 novice inspectors for the study and created two identical inspector groups. The inspectors used one of two playability heuristic sets to evaluate a mobile game. The second task was to analyze a predefined list of playability problems that were identified in the game. The inspectors' task was to annotate the problems with the proper playability heuristics. The selected playability heuristic sets were remarkably different, and hence, differences in the inspectors' ability to use them during the evaluation were presumed.

Hartson, Andre and Williges (2001) recommend that assessing usability evaluation methods (UEMs) for practical effectiveness should be done by measuring the thoroughness and validity of the method. These measures are based on a method's ability to identify usability problems in a product, and on how many of these problems are real problems that exist in the product. Sim, Read and Cockton (2009) have proposed that the validity of the heuristics should be evaluated using correctness and coverage as criteria. Correctness means that terminology and choice of words in the descriptions provide adequate information to the inspectors. Coverage indicates the extent of the domain knowledge that the heuristics have.

To estimate the thoroughness and validity of the usability evaluation method (UEM), we need to know the number of usability problems that exist in the design. In this study, we used a list that a playability expert generated. This is a recommended approach when we focus on actual problems that exist in a product (Hartson et al., 2001). As we are interested in the effects of a heuristic set on the novice inspectors' performance, we want to minimize the influence of the method or the evaluation procedure, and therefore, using the same method for creating an actual problem list should compensate for all the strengths and weaknesses of the method and provide a reliable estimation of the actual problems.

The inspectors should always evaluate the user interface and gameplay in playability evaluations. Further, the inspectors can evaluate other relevant aspects such as mobility in mobile games. User interface problems dominated in this study and both inspector groups reported approximately the same number of user interface problems. The difference in the evaluation coverage can be seen in the gameplay and mobility aspects. The inspectors who were using our playability heuristic set reported more playability problems related to the gameplay than the other inspector group. This difference was also statistically significant ($p=.008$). The group also reported playability problems related to mobility whereas the other group did not report any mobility problems.

The evaluation produced 374 playability problem reports which were further categorized into 64 unique problems. Both groups identified 42 out of 64 problems in the game. 17 problems were reported by at least nine inspectors (25% of the inspectors). These problems were inspected in

detail to see how the inspectors had annotated the reported problems, and to compare them to the heuristics that a playability expert had used. The correspondence between the selected heuristics from the playability expert and the novice inspectors was not on a very high level. With our playability heuristics, the inspectors achieved a 35% correspondence. In the other group, it was 33%. In the second task, both novice inspector groups annotated a predefined list of playability problems that were identified in the game by a playability expert. There were 44 playability problems in total and the problems were related to the main features of the game and they appeared in the gameplay, user interface, or mobility aspects. The correspondence of the selected heuristics improved compared to the first task. The group who used our playability heuristics achieved a 48% correspondence and for the other group it was 41%.

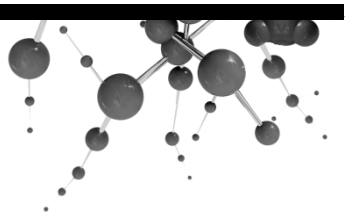
To assess the overall effectiveness of the playability heuristics, we calculated the thoroughness and the validity of the novice inspectors. The identified playability problems were compared to the standard problem list created by a playability expert. The overall effectiveness scores of the inspector groups were almost identical (0.235 and 0.231). However, Any-Two Agreement measurement for the group who used our playability heuristics was 19.6% and for the other group it was 12.3%, which indicates that there is a difference in internal consistency between the inspectors. Statistical analysis confirmed this and the difference was statistically significant ($p<.0001$).

After the evaluation tasks, the novice inspectors filled in a questionnaire about the use of the heuristic sets and estimated their correctness and coverage. Based on the questionnaire data, the inspectors seemed to be satisfied with the heuristic sets and statistical tests did not reveal significant differences between the inspector groups. This was a surprising result, since the data analysis revealed statistically significant differences in the evaluation results. This might be explained by the experiment setup. The inspectors did not see the other heuristic set and thus they did not have a reference point which would have allowed comparison of the heuristic sets. Moreover, it is possible that novice inspectors could not really judge their own performance during the evaluation and how the heuristic set had helped them in the evaluation task. This can be seen in the questionnaire responses and the evaluation data as some inspectors did not report any gameplay related problems, even though their response in the questionnaire indicates that the heuristic set helped them to evaluate the gameplay. Moreover, other novice inspectors and the playability expert identified playability problems in gameplay.

In summary, both playability heuristic sets can be used with the expert review method, but the inspector group who used our playability heuristics performed better in this study. Their evaluation was more

comprehensive and consistent compared to the other group. The categorization of the heuristics reflects the main aspects of playability which helps novice inspectors conduct the evaluation more efficiently.

The detailed description of the study has been presented in Papers VI and VII.



7 Conclusions

In this dissertation, I have presented a playability heuristic set that helps inspectors conduct mobile game evaluations by using the expert review method. The playability heuristics are grouped into five modules: *Game Usability*, *Gameplay*, *Mobility*, *Multi-Player,* and *Context-Aware.* The Game Usability and Gameplay modules are applicable to all kinds of games regardless of whether they are played on the PC, game consoles or mobile devices. They should be used in every game evaluation. The Game Usability module focuses on the user interface and the controls that the players use for interacting with a game. The Gameplay module contains core issues of the games related to game mechanics, story, goals and challenge. There are also heuristics for topics which influence the satisfaction of playing. For multi-player games, there is a separate module which deals with the social interaction of the players. The Mobility and the Context-Aware modules contain heuristics which are specific to games played on mobile devices and they take into account characteristics of the mobile devices and contextual factors which are more varying with mobile devices than other gaming platforms.

The playability heuristics are based on reviews of game design practices and literature, and hands on experience of evaluating mobile games and exploring what kinds of issues in design will cause problems for the players. Analytical work of defining playability heuristics continued throughout the dissertation and it refined the understandability of the heuristics to make them more complete and unambiguous. Each module went through a couple of iterations and they were used in several game evaluations. In addition, we collected feedback from external inspectors who used the playability heuristic set to evaluate mobile games.

As a result of defining playability heuristics we are also able to define what good playability means in video games. Our definition is the following:

> *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.*

This definition tells us the core issues of a playable game. It focuses on the two most critical game design aspects: the user interface and the gameplay. Naturally, there are also other factors influencing the game experience like the player's social interaction with other players in multi-player games and contextual factors in pervasive mobile games, but they were left out of the definition because it should be universally applicable to all kinds of games.

The validation of newly developed heuristics is an essential step in this kind of work. We have conducted several game evaluations during the development phase of the playability heuristics. Moreover, we have conducted comparison studies of domain-specific heuristics for video games to see the strengths and weaknesses of each heuristic set. The results indicate that our playability heuristic set helps inspectors to conduct the evaluation efficiently and the inspectors are able to discover playability problems that would have otherwise been left unnoticed. Another objective in the comparison studies was to gather data on whether the heuristic set is understandable to the inspectors. The results indicate that our heuristic set is able to describe discovered playability problems more accurately than other heuristic sets.

Finally, we conducted a comparison study of the expert review and the playtesting methods to see the usefulness of expert review in game evaluations. In game design literature, playtesting is the dominant method and other evaluation methods are usually not even mentioned. The results of our studies indicate that playability problems were identified quite consistently with these two methods. The expert review method accurately predicted the most serious playability problems in the games. Furthermore, expert review revealed playability problems related to the gameplay which were not identified in playtesting. Thus, it could be concluded that the method is a viable option for game evaluations and playability heuristics help the inspectors conduct the evaluation efficiently.

7.1 Reviewing the Contributions

The playability heuristic set is based on a review of game design literature which gives a solid foundation to the heuristic set. There are several

excellent game designers who have published their views on game design and good games in particular. Therefore I did not have to invent any of the game design aspects myself, but I could rely on the works done by others. I used my own knowledge and experience from the HCI domain to select which of these design principles are important from the playability point of view and how they should be formulated in the playability heuristic set.

The challenge in domain-specific heuristics for video games is to define heuristics which are suitable for all kinds of games. There are tens of game genres that have their own design characteristics. In addition, games are played on different gaming platforms which have unique input and output devices. The worst thing would be that the heuristics would require modifications every time they are used to evaluate the latest games.

During the dissertation work, we have evaluated games from multiple game genres that are played on different gaming platforms. The heuristic modules have not required any modifications between the studies once the final revision of the module was completed. This indicates that our heuristic set is robust and it can be used successfully in game evaluations.

In the recent studies of social games (outside of the scope of the dissertation work) it became apparent that one heuristic from the Game Usability module needs to be split up in order to better describe certain playability problems. This modification has been implemented in the Game Usability module and a new heuristic GU1b: "*A view to the gameworld supports smooth interaction and the camera behaves correctly*" was introduced. There was some indication of this already in the mobile game evaluation, but the appearance of the problem was not so frequent that it would have required action. Besides, the problem could be described with the original GU1 heuristic, but defining a new heuristic for the view and camera behavior will make it more visible. In these studies, we also discovered that social games which are played on social gaming platforms have specific playability problems which stem from the design characteristics of the social games and the viral marketing strategies that are used to attract a broader player population. It is still under consideration whether these games demand a new module to our heuristic set or whether they can be sufficiently covered by the current heuristics.

During the dissertation work, we put effort in the validation of the heuristics. The experimental research provided evidence of how the heuristics work in real game evaluations. We conducted evaluations which used different heuristic sets with the expert review method. In every study, our heuristic set was one of the two heuristic sets. Furthermore, we conducted game evaluations in which a game was evaluated using the expert review method and playtesting. All of these

experiments indicated that our heuristic set is capable of helping the inspector conduct the evaluation effectively.

7.2 THE LIMITATIONS OF THE STUDY

Mobile gaming is a complex domain and we have probably not yet fully seen the potential of mobile games that would utilize the characteristics of mobile devices and the contextual factors in the game design in a complete manner. Even today, most games, and especially the most popular games, played on mobile devices are similar to games that gamers play on stationary gaming devices.

During this dissertation work, we have mostly evaluated games that are traditional games which are played on mobile devices. Therefore, we also put a lot of effort into developing heuristics in the Game Usability and the Gameplay modules which cover playability issues in the user interface, interaction and game content. This is not at all a bad thing because those two modules are essential in any game evaluation and they are valid for all video games. Although the majority of the evaluated games resembled traditional games in this study, we were able to evaluate the mobility aspects and define playability heuristics which are important to mobile games, as well. An interesting notion from the playability heuristics in the Mobility module is that most of them are supplements to the heuristics in the Game Usability module. This indicates that mobile devices are similar to stationary gaming devices, but there are some issues which are slightly different when we consider the usability of the game and they should be made visible when evaluating mobile games. For that reason, they are included in a separate module rather than built into the Game Usability module.

The biggest difference in mobile gaming and playing games with stationary gaming devices is the context. The Context-Aware module contains heuristics that consider how context information as a game element influences the gaming experience and what kinds of playability problems might occur when the context impacts the game events. During the study, the availability of real mobile games was limited. We studied the influence of context on the gaming experience by using a mobile game which was designed and implemented in the Nokia Research Center. I was responsible for designing the user interface for the game. The Context-Aware module describes the main findings of these studies. In addition, we have used other published studies of pervasive games to explore important aspects of the playability of games that use context information as a game element.

Probably the Mobility and the Context-Aware modules are not yet as mature as the Game Usability and the Gameplay modules are because

only a few truly mobile games were studied. In the future, it would be fascinating to expand the current work and study various mobile games that utilize the features of mobile devices (e.g. sensors, camera, and GPS) in the game design to see how they influence playability and the gaming experience and further playability heuristics.

7.3 Implications for Practitioners

The playability heuristics presented in this dissertation are ready to be used by practitioners in the game industry. Even though the heuristics are intended to be used during the evaluation task to find out possible playability problems, they are also useful during the design phase of a game. Game designers can use them to avoid common pitfalls in design. It has been said that to be fair to designers, game design should be evaluated based on the same criteria that have been used for designing the game. Appendix A provides a consolidated list of heuristics that inspectors and game designers can use in their daily work.

The modular structure makes playability heuristics very flexible in game development projects. The inspectors or designers can use one or more modules during the task according to their needs. The smaller number of heuristics is easier to cover and each module is targeted for a specific purpose. For example, if a person is designing the user interface of the shell menu or the gameplay, he needs to work with the game usability heuristics. Correspondingly, in multi-player games, an inspector needs to consider multi-player heuristics when evaluating the social interaction of the players.

The purpose of the game evaluations with the expert review method is to remove design mistakes that would cause problems to the players. Expert review is one type of activity which tries to ensure the quality of a game and it has a different role compared to playtesting and quality assurance. Expert review is a versatile and flexible method to explore the game design frequently and give quick feedback to the designers. It can be used in any phase of the design as inspectors are able to position and focus themselves to the readiness level of the design. The method is most advantageous when it is used before playtesting and quality assurance activities take place.

The dissemination of the heuristics has already started and there are several scientific articles available in which our playability heuristics are cited. Furthermore, the playability heuristics have been referred to at least in four books (Laitinen, 2008; Markopoulos, Read, MacFarlane, & Höysniemi, 2008; Shneiderman & Plaisant, 2009; Hochleitner, Hochleitner, Graf, & Tscheligi, 2015) about the expert review method and game evaluations. This gives some indication that our playability heuristics are

considered relevant and they are recommended to be used both in scientific studies and the industry.

7.4 Final Remarks

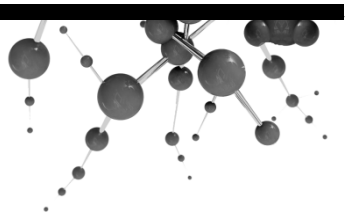
In this dissertation, the primary focus was to identify what constitutes the playability of mobile games or video games in general and how to define domain-specific heuristics for an evaluation task. Playability is a multifaceted structure which has to deal with the actual content of a game in addition to the user interface. In many cases, the evaluation also needs to consider the social interaction of the players because in many games multiple players are acting in the same gameworld and influencing each other. This is a clear difference compared to traditional usability evaluations of products because they often only focus on the user interface. This study provides insights on how to take this diversity into account and what is required to evaluate video games comprehensively.

Recently, we have explored the applicability of playability heuristics for more domain-specific games and especially for social network games which are played on social network platforms such as Facebook. Although the playability heuristics presented in this thesis are capable of discovering many playability problems in social games (Paavilainen, Korhonen, & Alha, 2014), there are also some domain-specific problems which emerge from the design characteristics of these games (Paavilainen, Alha, & Korhonen, 2015). Defining playability heuristics for these problems was out of the scope of this work.

From the evaluation methodology point of view, game evaluations provide interesting topics of research. As video games are very engaging and gamers are often immersed in the gameworld, traditional evaluation methods may not work in an optimal way. In this study, we noticed that even with the expert review method, the procedure of conducting an evaluation is slightly different than in productivity software evaluations. More research on evaluation methodology is required to understand the challenges of game evaluations. One possible approach could be to separate the playing and the evaluating tasks. Evaluation methods could be directed towards post-game analysis where a skillful player from the target player group plays the game and the inspectors use video and other recordings of the play session to measure the game, the player, and the interaction with game elements and other players to determine possible playability problems in the design and to see the player's reactions and responses to the game content. Capturing the intensity of gameplay and the effect of the player's skills and expertise are hard to include in the evaluation, but on the other hand, this is a challenge when evaluations also include actual content and not just the user interface elements. The

playability heuristics in this approach could help the inspectors to pay attention to the right aspects of the game design.

The game evaluation methodology is not complete yet and hopefully this research inspires other game researchers to continue an evaluation methodology research branch where new methods are developed specifically for game evaluations.



8 References

Adams, E. & Rollings, A. (2007). *Game Design and Development: Fundamentals of Game Design*. Berkeley, CA, USA: Pearson Prentice Hall.

Anttila, A. & Jung, Y. (2006). Discovering Design Drivers for Mobile Media Solutions. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA '06, Montreal, Quebec, Canada)*, 219–224. New York, NY, USA: ACM. doi: 10.1145/1125451.1125497

Ardito, C., Costabile, M. F., De Marsico, M., Lanzilotti, R., Levialdi, S., Plantamura, P., Roselli, T., Rossano, V. & Tersigni, M. (2004). *Towards Guidelines for Usability of e-Learning Applications*. In Stary, C. & Stephanidis, C. (Eds.), User-Centered Interaction Paradigms for Universal Access in the Information Society (8th ERCIM Workshop on User Interfaces for All, (UI4All)), LNCS 3196, 185-202. Berlin Heidelberg, Germany: Springer-Verlag. doi: 10.1007/978-3-540-30111-0_16

Baker, K., Greenberg, S. & Gutwin, C. (2001). Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. In *Proceedings of the 8th IFIP Working Conference on Engineering for Human-Computer Interaction (EHCI'01)*, 123-139. doi: 10.1007/3-540-45348-2_14

Baker, K., Greenberg, S. & Gutwin, C. (2002). Empirical Development of a Heuristic Evaluation Methodology for Shared Workspace Groupware. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)*, 96–105. New York, NY, USA: ACM Press. doi: 10.1145/587078.587093

Ballagas, R. A., Kratz, S. G., Borchers, J., Yu, E., Walz, S. P., Fuhr, C. O., Hovestadt, L. & Tann, M. (2007). REXplorer: A Mobile, Pervasive Spell-Casting Game for Tourists. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA '07, San Jose, CA, USA)*, 1929–1934. New York, NY, USA: ACM Press. doi: 10.1145/1240866.1240927

Barry, I. (2005). Game Design. Chapter 2 In Rabin, S. (Ed.), *Introduction to Game Development*. 99–160. Massachusetts, MA, USA: Charles River Media.

Bartle, R. C. (1997). *Hearts, Clubs, Diamonds, Spades: Player who suit MUD*. http://mud.co.uk/richard/hcds.htm (Accessed May 13, 2014)

Baughman, N. E. & Levine, B. N. (2001). Cheat-Proof Playout for Centralized and Distributed Online Games. In *Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 104–113. New York, NY, USA: IEEE. doi: 10.1109/INFCOM.2001.916692

Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Duncan, R., Benford, S., Capra, M. & Hampshire, A. (2006). Interweaving Mobile Games with Everyday Life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*, 417–426. New York, NY, USA: ACM Press. doi: 10.1145/1124772.1124835

Benford, S., Flintham, M., Drozd, A., Anastasi, R., Rowland, D., Tandavanitj, N., Adams, M., Row-Farr, J., Oldroyd, A. & Sutton, J. (2004). Uncle Roy All Around You: Implicating the City in a Location-Based Performance. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'04)*, New York, NY, USA: ACM.

Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E., Loh, S. & Reeves, T. C. (2002). *Usability and Instructional Design Heuristics for E-Learning Evaluation*. In Barker, P. & Rebelsky, S. (Eds.), EdMedia: World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002 (EdMedia 2002), 1615-1621. Norfolk, VA, USA: Association for the Advancement of Computing in Education (AACE).

Benyon, D., Turner, P. & Turner, S. (2005). *Designing Interactive Systems: People, Activities; Context, Technologies*. Essex, UK: Pearson Education Limited.

Bernhaupt, R. (2010). *Evaluating User Experience in Games: Concepts and Methods*. London, UK: Springer-Verlag.

Bernhaupt, R., Boldt, A., Mirlacher, T., Wilfinger, D. & Tscheligi, M. (2007). Using Emotion in Games: Emotional Flower. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'07)*, 41–48. New York, NY, USA: ACM Press. doi: 10.1145/1255047.1255056

Berry, B. (2003). Adapting Heuristics for Notification Systems. In *Proceedings of the ACM Southeast Conference*, 144–149. New York, NY, USA: ACM Press.

Bertini, E., Catarci, T., Dix, A., Gabrielli, S., Kimani, S. & Santucci, G. (2009). Appropriating Heuristic Evaluation for Mobile Computing. *International Journal of Mobile Human Computer Interaction*, *1* (1), 20–41. doi: 10.4018/jmhci.2009010102

Bias, R. & Mayhew, D. (2005). *Cost-Justifying Usability: An Update for the Internet Age, Second Edition*. San Francisco, CA, USA: Morgan Kaufmann Publisher.

Bickford, P. (1997). *Interface Design: The Art of Developing Easy-to-Use Software*. Chessnut Hill, MA, USA: Academic Press.

Birk, M. & Mandryk, R. L. (2013). Control Your Game-Self: Effects of Controller Type on Enjoyment, Motivation, and Personality in Game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*, 685-694. New York, NY, USA: ACM. doi: 10.1145/2470654.2470752

Björk, S., Falk, J., Hansson, R. & Ljungstrand, P. (2001). Pirates! Using the Physical World as a Game Board. In *Proceedings of the IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT)*, 423–430. Amsterdam, The Netherlands: IOS Press.

Björk, S. & Holopainen, J. (2005). *Patterns in Game Design*. Hingham, MA, USA: Charles River Media.

Bond, M. & Beale, R. (2009). What Makes a Good Game? Using Reviews to Inform Design. In *Proceedings of the British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (BSC-HCI'09)*, 418–422. Swinton, UK: British Computer Society.

Breslin, S. (2009). *The History and Theory of Sandbox Gameplay*. http://www.gamasutra.com/view/feature/132470/the_history_and_theory_of_sandbox_.php?print=1 (Accessed December 12, 2013)

Broll, G. & Benford, S. (2005). Seamful Design for Location-Based Mobile Games. In *Proceedings of the International Conference on Entertainment*

Computing (ICEC 2005), 155–166. Berlin Heidelberg: Springer. doi: 10.1007/11558651_16

Bruckman, A., Curtis, P., Figallo, C. & Laurel, B. (1994). Approaches to Managing Deviant Behavior in Virtual Communities. In *Proceedings of the Conference Companion on Human Factors in Computing Systems (CHI '94)*, 183–184. New York, NY, USA: ACM. doi: 10.1145/259963.260231

Cairns, P., Li, J., Wang, W. & Nordin, A. I. (2014). The Influence of Controllers on Immersion in Mobile Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*, 371-380. New York, NY, USA: ACM. doi: 10.1145/2556288.2557345

Carroll, J. M. (2004). Beyond Fun. *interactions*, *11* (5), 38–40. doi: 10.1145/1015530.1015547

Carroll, J. M. & Rosson, M. B. (1992). Getting Around the Task-Artifact Cycle: How to Make Claims and Design by Scenario. *ACM Transactions on Information Systems*, *10* (2), 181–212. doi: 10.1145/146802.146834

Chandler, A. & Finney, J. (2005). On the Effects of Loose Causal Consistency in Mobile Multiplayer Games. In *Proceedings of the ACM SIGCOMM workshop on Network and system support for Games (NetGames '05)*, 1–11. New York, NY, USA: ACM. doi: 10.1145/1103599.1103602

Chattratichart, J. & Brodie, J. (2002). Extending the Heuristic Evaluation Method through Contextualisation. In *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES)*, 641–645. Thousand Oaks, CA, USA: SAGE Publications.

Chen, K.-T., Huang, C.-Y., Huang, P. & Lei, C.-L. (2006). An Empirical Evaluation of TCP Performance in Online Games. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'06)*, Article No. 5. New York, NY, USA: ACM. doi: 10.1145/1178823.1178830

Chen, V. H.-h. & Duh, H. B.-L. (2007). Understanding Social Interaction in World of Warcraft. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'07)*, 21–24. New York, NY, USA: ACM. doi: 10.1145/1255047.1255052

Cherny, L., Clanton, C. & Ostrom, E. (1997). Entertainment is a Human Factor: A CHI97 Workshop on Game Design and HCI. *SIGCHI Bulletin*, *29* (4),

Cheung, G. K., Zimmermann, T. & Nagappan, N. (2014). The First Hour Experience: How the Initial Play Can Engage (or Lose) New Players. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '14)*, 57-66. New York, NY, USA: ACM Press. doi: 10.1145/2658537.2658540

Chilana, P. K., Wobbrock, J. O. & Ko, A. J. (2010). Understanding Usability Practices in Complex Domains. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'10)*, 2337–2346. New York, NY, USA: ACM Press. doi: 10.1145/1753326.1753678

Clanton, C. (1998). An Interpreted Demonstration of Computer Game Design. In *Proceedings of the SIGCHI Conference Summary on Human Factors in Computing Systems (CHI'98)*, 1–2. New York, NY, USA: ACM Press. doi: 10.1145/286498.286499

Clanton, C. (2000). Lessons from Game Design. Chapter 10 In Bergman, E. (Ed.), *Information Appliances and Beyond: Interaction Design for Consumer Products*. 299–334. San Diego, CA, USA: Morgan Kaufmann.

Clarke, L. (1991). The Use of Scenarios by User Interface Designers. In *Proceedings of the People and Computers VI (BCS-HCI)*, 103–115. New York, NY, USA: Cambridge University Press.

Clinton, K. (2010). *Agile Game Development With Scrum: Teams*. http://www.gamasutra.com/view/feature/134412/agile_game_development_with_scrum_.php (Accessed December 12, 2013)

Cockton, G. & Woolrych, A. (2002). Sale Must End: Should Discount Methods Be Cleared Off HCI's Shelves? *Interactions*, *9* (5), 13–18. doi: 10.1145/566981.566990

Cornett, S. (2004). The Usability of Massively Multiplayer Online Roleplaying Games: Designing for New Users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04)*, 703–710. New York, NY, USA: ACM Press. doi: 10.1145/985692.985781

Courage, C. & Baxter, K. (2005). *Understanding Your Users: A Practical Guide to User Requirements Methods, Tools and Techniques*. San Francisco, CA, USA: Morgan Kaufmann.

Crabtree, A., Benford, S., Capra, M., Flintham, M., Drozd, A., Tandavanitj, N., Adams, M. & Row Farr, J. (2007). The Cooperative Work of Gaming: Orchestrating a Mobile SMS Game. *Computer Supported*

Cooperative Work (CSCW), *16* (1-2), 167–198. doi: 10.1007/s10606-007-9048-1

Cronholm, S. & Bruno, V. (2008). Do You Need General Principles or Concrete Heuristics?: A Model for Categorizing Usability Criteria. In *Proceedings of the Australasian Conference on Computer-Human Interaction (OZCHI)*, 105–111. New York, NY, USA: ACM Press. doi: 10.1145/1517744.1517779

Csikszentmihalyi, M. (1991). *Flow: The Psychology of Optimal Experience*. New York, USA: Harper Perennial Modern Classics.

Cui, Y., Chipchase, J. & Jung, Y. (2007). *Personal TV: A Qualitative Study of Mobile TV Users*. In Cesar, P., Chorianopoulos, K. & Jensen, J. (Eds.), Interactive TV: a Shared Experience (EUROITV 2007), LNCS 4471, 195–204. Berlin Heidelberg, Germany: Springer. doi: 10.1007/978-3-540-72559-6_21

Davidsson, O., Peitz, J. & Björk, S. (2004). *Game Design Patterns for Mobile Game Design*. Nokia Research Center.

Desmet, P. & Hekkert, P. (2007). Framework of Product Experience. *International Journal of Design*, *1* (1), 57–66. http://www.ijdesign.org/ojs/index.php/IJDesign/article/view/66/15

Desurvire, H., Caplan, M. & Toth, J. A. (2004). Using Heuristics to Evaluate the Playability of Games. In *Extended Abstracts on SIGCHI Conference on Human Factors in Computing Systems (CHI EA '04, Vienna, Austria)*, 1509–1512. New York, NY, USA: ACM Press. doi: 10.1145/985921.986102

Desurvire, H., Kondziela, J. & Atwood, M. E. (1992). What is Gained and Lost When Using Methods Other Than Empirical Testing. In *Proceedings of the Conference on People and Computers VII (BCS-HCI)*, 89–102. New York, NY, USA: Cambridge University Press. doi: 10.1145/1125021.1125115

Desurvire, H. & Wiberg, C. (2008). Master of the Game: Assessing Approachability in Future Game Design. In *Extended Abstracts on SIGCHI Conference on Human Factors in Computing Systems (CHI EA '08, Florence, Italy)*, 3177-3182. New York, NY, USA: ACM. doi: 10.1145/1358628.1358827

Desurvire, H. & Wiberg, C. (2009). *Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration*. In Ozok, A. A. & Zaphiris, P. (Eds.), Third International Conference Online Communities and Social Computing, OCSC 2009, Held as Part of

HCI International 2009 (HCI-I), LNCS 5621, 557–566. Berlin Heidelberg, Germany: Springer. doi: 10.1007/978-3-642-02774-1_60

Dick, M., Wellnitz, O. & Wolf, L. (2005). Analysis of Factors Affecting Players' Performance and Perception in Multiplayer Games. In *Proceedings of the ACM SIGCOMM workshop on Network and system support for games (NetGames '05)*, 1–7. New York, NY, USA: ACM. doi: 10.1145/1103599.1103624

Dixon, H., Mitchell, V. A. & Harker, S. D. P. (2004). Mobile Phone Games: Understanding the User Experience. In McDonagh, D., Hekkert, P., van Erp, J. & Gyi, D. (Eds.), *Design and Emotion: The Experiences of Everyday Things*. 256–261. London, UK: Taylor and Francis.

Drury, J. (2001). Developing Heuristics for Synchronous Collaborative Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI EA'01)*, 447-448. ACM. doi: 10.1145/634067.634326

Ducheneaut, N. & Moore, R. J. (2004). The Social Side of Gaming: A Study of Interaction Patterns in a Massively Multiplayer Online Game. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'04)*, 360–369. New York, NY, USA: ACM. doi: 10.1145/1031607.1031667

Ducheneaut, N., Yee, N., Nickell, E. & Moore, R. J. (2006). "Alone together?": Exploring the Social Dynamics of Massively Multiplayer Online Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*, 407–416. New York, NY, USA: ACM. doi: 10.1145/1124772.1124834

Ducheneaut, N., Yee, N., Nickell, E. & Moore, R. J. (2007). The Life and Death of Online Gaming Communities: A Look at Guilds in World of Warcraft. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*, 839–848. New York, NY; USA: ACM. doi: 10.1145/1240624.1240750

Dumas, J. S. & Loring, B. (2008). *Moderating Usability Tests: Principles & Practices for Interacting*. Burlington, MA, USA: Morgan Kaufmann.

Dumas, J. S., Molich, R. & Jeffries, R. (2004). Describing Usability Problems: Are We Sending the Right Message? *interactions*, *11* (4), 24-29. doi: 10.1145/1005261.1005274

Dumas, J. S. & Redish, J. C. (1999). *A Practical Guide to Usability Testing*. Portland, OR, USA: intellect Books.

Dykstra, D. J. (1993). *A Comparison of Heuristic Evaluation and Usability Testing: The Efficacy of a Domain-specific Heuristic Checklist*. PhD Dissertation in Texas A & M University.

Egenfeldt-Nielsen, S., Smith, J. H. & Pajares Tosca, S. (2008). *Understanding Video Games: The Essential Introduction*. New York, NY, USA: Routledge.

El-Nasr, M. S., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H. & Mah, S. (2010). Understanding and Evaluating Cooperative Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'10)*, 253-262. New York, NY, USA: ACM. doi: 10.1145/1753326.1753363

Ermi, L. & Mäyrä, F. (2005). Fundamental Components of the Gameplay Experience: Analysing Immersion. In *Proceedings of the DIGRA 2005 Conference: Changing Views - Worlds in Play*, Vancouver, BC, Canada: University of Vancouver.

Evans, C. & Sabry, K. (2003). Evaluation of the Interactivity of Web-based Learning Systems: Principles and Process. *Innovations in Education and Teaching International*, *40* (1), 89-99. doi: 10.1080/1355800032000038787

Fabricatore, C., Nussbaum, M. & Rosas, R. (2002). Playability in Action Videogames: A Qualitative Design Model. *Human-Computer Interaction*, *17* (4), 311–368. doi: 10.1207/S15327051HCI1704_1

Falstein, N. (2003). Better by Design: Good Points Don't Go Down. *Game Developer*, *10* (2), 28.

Falstein, N. (2005). Understanding Fun - The Theory of Natural Funativity. Chapter 2.1 In Rabin, S. (Ed.), *Introduction to Game Development*. 71–97. Hingham, Massachutetts, USA: Charles River Media, Inc.

Falstein, N. & Barrywood, H. (2006). *The 400 Project*. http://www.theinspiracy.com/the-400-project.html (Accessed April 5, 2013)

Farrell, G. & Farrell, V. (2012). Application of Domain Specific Heuristics to an Innovative Computer Based Assessment Strategy. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, 123-129. New York, NY, USA: ACM. doi: 10.1145/2414536.2414557

Federoff, M. A. (2002). *Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games*. Master Thesis in Department of Telecommunications, Indiana University.

Foo, C. Y. & Koivisto, E., M. I. (2004). Defining Grief Play in MMORPGs: Player and Developer Perceptions. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'04)*, 245–250. New York, NY, USA: ACM. doi: 10.1145/1067343.1067375

Fox, B. (2005). *Game Interface Design*. Boston, MA, USA: Thompson Course Technology.

Fridja, N. H. (1986). *The Emotions*. Cambridge, UK: Cambridge University Press.

Friedl, M. (2002). *Online Game Interactivity Theory*. Boston, MA, USA: Charles River Media, Inc.

Fritsch, T., Ritter, H. & Schiller, J. (2005). The Effect of Latency and Network Limitations on MMORPGs: A Field Study of Everquest2. In *Proceedings of the ACM SIGCOMM Workshop on Network and System Support for Games (NetGames'05)*, 1–9. New York, NY, USA: ACM. doi: 10.1145/1103599.1103623

Fullerton, T., Swain, C. & Hoffman, S. (2004). *Game Design Workshop: Designing, Prototyping, and Playtesting Games*. San Francisco, CA, USA: CMP Books.

Fullerton, T., Swain, C. & Hoffman, S. S. (2008). *Game Design Workshop: A Playcentric Approach to Creating Innovative Games*. Burlington, MA, USA: Morgan Kaufmann.

Følstad, A., Anda, B. C. D. & Sjøberg, D. I. K. (2010). The Usability Inspection Performance of Work-Domain Experts: An Empirical Study. *Interacting with Computers*, *22* (2), 75–87. doi: 10.1016/j.intcom.2009.09.001

Garris, R., Ahlers, R. & Driskell, J. E. (2002). Games, Motivation, and Learning: A Research and Practice Model. *Simulation Gaming*, *33* (4), 441–467. doi: 10.1177/1046878102238607

Gautier, L. & Diot, C. (1998). Design and Evaluation of MiMaze, a Multi-Player Game on the Internet. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, 233–236. New York, NY, USA: IEEE. doi: 10.1109/MMCS.1998.693647

Geerts, D. & De Grooff, D. (2009). Supporting the Social Uses of Television: Sociability Heuristics for Social TV. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, 595-604. ACM. doi: 10.1145/1518701.1518793

Gilbertson, P., Coulton, P., Chehimi, F. & Vajk, T. (2008). Using "Tilt" as an Interface to Control "No-Button" 3-D Mobile Games. *Computers in Entertainment*, *6* (3), Article No. 38. doi: 10.1145/1394021.1394031

Gray, W. D. & Salzman, M. C. (1998). Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. *Human-Computer Interaction*, *13* (3), 203–261. doi: 10.1207/s15327051hci1303_2

Hakkarainen, M. & Woodward, C. (2005). SymBall: Camera Driven Table Tennis for Mobile Phones. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'05)*, 391–392. New York, NY, USA: ACM Press. doi: 10.1145/1178477.1178565

Hall, J. (2005). Future of Games: Mobile Gaming. Chapter 3 In Raessens, J. & Goldstein, J. (Eds.), *Handbook of Computer Game Studies*. 47–55. Cambridge, Massachusetts, USA: MIT Press.

Hallford, N. & Hallford, J. (2001). *Swords and Circuitry: A Designer's Guide to Computer Role Playing Games*. Roseville, CA, USA: Prima Publishing.

Hara, M. & Ovaska, S. (2014). Heuristics for Motion-Based Control in Games. In *Proceedings of the Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI'14)*, 697-706. New York, NY, USA: ACM. doi: 10.1145/2639189.2639246

Hartson, H. R., Andre, T. S. & Williges, R. C. (2001). Criteria For Evaluating Usability Evaluation Methods. *International Journal of Human-Computer Interaction*, *13* (4), 373 - 410. doi: 10.1207/S15327590IJHC1501_13

Hassenzahl, M. (2003). The Thing and I: Understanding the Relationship between User and Product. Chapter 3 In Blythe, M., Monk, A. F., Overbeeke, K. & Wright, P. (Eds.), *Funology: From Usability to Enjoyment*. 31–42. Dordrecht, The Netherlands: Kluwer Academic Publisher.

Heikkinen, K., Strömberg, H., Koivisto, E. M. I. & Suomela, R. (2006). Heuristic Rule-set for Mobile Context Awareness. In *Proceedings of the IMAC Workshop in conjuction with MobileHCI 2006*, http://hci.icts.sbg.ac.at/context/imac

Henricksen, K. & Indulska, J. (2004). Modelling and Using Imperfect Context Information. In *Proceedings of the IEEE Annual Conference on Pervasive Computing and Communications Workshops (PERCOMW'04)*,

33–37. Washington, DC, USA: IEEE. doi: 10.1109/PERCOMW.2004.1276901

Hertzum, M. & Jacobsen, N. E. (2001). The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods. *International Journal of Human-Computer Interaction*, *13* (4), 421–443. doi: 10.1207/S15327590IJHC1304_05

Hinske, S., Langheinrich, M. & Lampe, M. (2008). Towards Guidelines for Designing Augmented Toy Environments. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS'08)*, 78–87. New York, NY, USA: ACM Press. doi: 10.1145/1394445.1394454

Hochleitner, C., Hochleitner, W., Graf, C. & Tscheligi, M. (2015). A Heuristic Framework for Evaluating User Experience in Games. Chapter 9 In Bernhaupt, R. (Ed.), *Game User Experience Evaluation*. 187-206. Cham, Switzerland: Springer International Publishing.

Huizinga, J. (1955). *Homo Ludens; a Study of the Play-Element in Culture*. Boston, MA, USA: Beacon Press.

Hunicke, R., LeBlanc, M. & Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. In *Proceedings of the Game Design and Tuning Workshop at the Game Developers Conference*, http://www.aaai.org/Papers/Workshops/2004/WS-04-04/WS04-04-001.pdf

Hvannberg, E. T., Law, E. L.-C. & Lárusdóttir, M. K. (2007). Heuristic Evaluation: Comparing Ways of Finding and Reporting Usability Problems. *Interacting with Computers*, *19* (2), 225–240. doi: 10.1016/j.intcom.2006.10.001

Inostroza, R., Rusu, C., Roncagliolo, S., Jimenez, C. & Rusu, V. (2012). Usability Heuristics for Touchscreen-based Mobile Devices. In *Extended Abstracts on Ninth International Conference on Information Technology: New Generations (ITNG, Las Vegas, NV, USA)*, 662-667. New York, NY, USA: IEEE. doi: 10.1109/itng.2012.134

Isbister, K. & Schaffer, N. (2008). *Game Usability: Advice from the Experts for Advancing the Player Experience*. Burlington, MA, USA: Morgan Kaufmann.

ISO (1998). ISO 9241-11: *Ergonomic Requirements for Office Work with Visual Display Terminals (VDT)s - Part 11: Guidance on Usability*. International Organization for Standardization, Geneva, Switzerland.

ISO (2010). ISO 9241-210: *Ergonomics of Human-System Interaction - Part 210: Human Centered Design Processes for Interactive Systems*. International Organization for Standardization, Geneva, Switzerland.

Jacko, J. A. & Salvendy, G. (1996). Hiearchical Menu Design: Breadth, Depth, and Task Complexity. *Perceptual and Motor Skills*, *82* (3c), 1187–1201. doi: 10.2466/pms.1996.82.3c.1187

Jacobsen, N. E., Hertzum, M. & John, B. E. (1998). The Evaluator Effect in Usability Tests. In *Proceedings of the SIGCHI Conference Summary on Human Factors in Computing Systems (CHI '98)*, 255–256. New York, NY, USA: ACM Press. doi: 10.1145/286498.286737

Jaferian, P., Hawkey, K., Sotirakopoulos, A., Velez-Rojas, M. & Beznosov, K. (2014). Heuristics for Evaluating IT Security Management Tools. *Human–Computer Interaction*, *29* (4), 311-350. doi: 10.1080/07370024.2013.819198

Jeffries, R., Miller, J. R., Wharton, C. & Kathy, U. (1991). User Interface Evaluation in the Real World: A Comparison of Four Techniques. In *Proceedings of the Conference on Human Factors on Computing Systems (CHI '91)*, 119–124. New York, NY, USA: ACM Press. doi: 10.1145/108844.108862

Jegers, K. (2008). Investigating the Applicability of Usability and Playability Heuristics for Evaluation of Pervasive Games. In *Proceedings of the International Conference on Internet and Web Applications and Services (ICIW)*, 656–661. Washington, DC, USA: IEEE. doi: 10.1109/ICIW.2008.54

Jenkins, R., Lee, M. & Archambault, R. (2010). *The Influence of Professional Critic Reviews*. Report in EEDAR/SMU Behavioral Study, Southern Methodist University, Guildhall, Electronic Entertainment Design and Research.

Jensen, C., Davis, J. & Farnham, S. (2002). Finding Others Online: Reputation Systems for Social Online Spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'02)*, 447–454. New York, NY, USA: ACM. doi: 10.1145/503376.503456

Johnson, D. & Wiles, J. (2003). Effective Affective User Interface Design in Games. *Ergonomics*, *46* (13-14), 1332–1345. doi: 10.1080/00140130310001610865

Jones, M. & Marsden, G. (2006). *Mobile Interaction Design*. Chichester, UK: John Wiley & Sons.

Jordan, P. W. (2000). *Designing Pleasurable Products*. Florida, USA: CRC Press.

Jumisko-Pyykkö, S. & Vainio, T. (2010). Framing the Context of Use for Mobile HCI. *International Journal of Mobile Human Computer Interaction*, *2* (4), 1–18. doi: 10.4018/jmhci.2010100101

Juul, J. (2005). *Half-Real: Video Games between Real Rules and Fictional Worlds*. Cambridge, MA, USA: MIT Press.

Järvinen, A. (2002a). *Gran Stylissimo: The Audiovisual Elements and Styles in Computer and Video Games*. In Mäyrä, F. (Ed.), Computer Games and Digital Cultures (CGDC), 113–128. Tampere, Finland: Tampere University Press.

Järvinen, A. (2002b). *Milloin pelistä tulee mobiili?* http://www.m-cult.net/mediumi/article.html?articleId=29&print=1&lang=fi (Accessed November, 20, 2008)

Järvinen, A., Heliö, S. & Mäyrä, F. (2002). *Communication and Community in Digital Entertainment Services: Prestudy Research Report*. Report in Hypermedia Laboratory Net Series, Report No. 2 University of Tampere.

Karat, C.-M., Campbell, R. & Fiegel, T. (1992). Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'92)*, 397-404. New York, NY, USA: ACM Press. doi: 10.1145/142750.142873

Karvonen, J. (2005). *Mobiilipelin Pelattavuuden Arviointi*. Master's Thesis in Department of Computer Science, University of Jyväskylä. http://urn.fi/URN:NBN:fi:jyu-2006151

Kersey, I. (2013). *Game Developers: To Hell With Consoles, Hello Mobile!* http://www.intomobile.com/2013/03/01/mobile-platforms-gain-popularity-developers/ (Accessed October, 16, 2013)

Kiili, K. (2005). *On Educational Game Design: Building Blocks of Flow Experience*. Ph.D. Thesis in Tampere University of Technology - Pori.

Kim, K., Jacko, J. A. & Salvendy, G. (2011). Menu Design for Computers and Cell Phones: Review and Reappraisal. *International Journal of Human-Computer Interaction*, *27* (4), 383–404. doi: 10.1080/10447318.2011.540493

Kirschner, D. & Williams, P. J. (2013). Experts and Novices or Expertise? Positioning Players through Gameplay Reviews. In *Proceedings of*

the DiGRA International Conference: DeFragging Game Studies (DiGRA'13), DIGRA.

Koivisto, E. M. I. (2003). Supporting Communities in Massively Multiplayer Online Role-Playing Games by Game Design. In *Proceedings of the DIGRA 2003 Conference: Level Up*, Utrecht, The Netherlands: University of Utrecht.

Korhonen, H. (2011). The Explanatory Power of Playability Heuristics. In *Proceedings of the Advances in Computer Entertainment Technology (ACE'11)*, Article No. 40. New York, NY, USA: ACM. doi: 10.1145/2071423.2071473

Korhonen, H., Arrasvuori, J. & Väänänen-Vainio-Mattila, K. (2010). Analysing User Experience of Personal Mobile Products through Contextual Factors. In *Proceedings of the International Conference Mobile and Ubiquitous Multimedia (MUM 2010)*, 1–10. New York, NY, USA: ACM. doi: 10.1145/1899475.1899486

Korhonen, H. & Koivisto, E. M. I. (2006). Playability Heuristics for Mobile Games. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'06)*, 9–16. New York, NY, USA: ACM Press. doi: 10.1145/1152215.1152218

Korhonen, H. & Koivisto, E. M. I. (2007). Playability Heuristics for Mobile Multi-Player Games. In *Proceedings of the International Conference on Digital Interactive Media in Entertainment and Arts (DIMEA 2007)*, 28-35. New York, NY, USA: ACM Press. doi: 10.1145/1306813.1306828

Korhonen, H., Paavilainen, J. & Saarenpää, H. (2009). Expert Review Method in Game Evaluations – Comparison of Two Playability Heuristic Sets. In *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era (MindTrek '09)*, 74-81. New York, NY, USA: ACM. doi: 10.1145/1621841.1621856

Korhonen, H., Saarenpää, H. & Paavilainen, J. (2008). *Pervasive Mobile Games – A New Mindset for Players and Developers*. In Markopoulos, P., de Ruyter, B., Ijsselsteijn, W. & Rowland, D. (Eds.), Fun and Games (Fun'n'Games 2008), LNCS 5294, 21-32. Berlin Heidelberg, Germany: Springer-Verlag. doi: 10.1007/978-3-540-88322-7_3

Kultima, A. (2009). Casual Game Design Values. In *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era (MindTrek '09)*, 58–65. New York, NY, USA: ACM. doi: 10.1145/1621841.1621854

Kultima, A. & Stenros, J. (2010). Designing Games for Everyone: The Expanded Game Experience Model. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology (Futureplay '10)*, 66–73. New York, NY, USA: ACM. doi: 10.1145/1920778.1920788

Kuparinen, L., Silvennoinen, J. & Isomäki, H. (2013). Introducing Usability Heuristics for Mobile Map Applications. In *Proceedings of the 26th International Cartographic Conference (ICC '13)*, International Cartographic Association.

Kurosu, M., Matsuura, S. & Sugizaki, M. (1997). Categorical Inspection Method - Structured Heuristic Evaluation (sHEM). In *Proceedings of the IEEE International Conference on Systems, Man, and Cypernetics (volume: 3) (ICSMC)*, 2613–2618. Washington, DC, USA: IEEE. doi: 10.1109/ICSMC.1997.635329

Kücklich, J. (2004). *Play and Playability as Key Concepts in New Media Studies*. White Paper. Available at: http://www.playability.de/play.pdf (Accessed November 11, 2011)

Köffel, C. & Haller, M. (2008). Heuristics for the Evaluation of Tabletop Games. In *Proceedings of the Workshop on Evaluating User Experiences in Games at CHI2008*, http://workshops.icts.sbg.ac.at/chi2008uxgames/

Laitinen, S. (2006). Do Usability Expert Evaluation and Test Provide Novel and Useful Data for Game Development? *Journal of Usability Studies*, *1* (2), 64–75.

Laitinen, S. (2008). Usability and Playability Expert Evaluation. Chapter 7 In Isbister, K. & Schaffer, N. (Eds.), *Game Usability - Advice from the Experts for Advancing the Player Experience*. 91–111. Burlington, MA, USA: Morgan-Kaufmann.

Lavery, D., Cockton, G. & Atkinson, M. P. (1997). Comparison of Evaluation Methods Using Structured Usability Problem Reports. *Behaviour & Information Technology*, *16* (4), 246–266. doi: 10.1080/014492997119824

LeBlanc, M. (2006). Tools for Creating Dramatic Game Dynamics. In Salen, K. & Zimmerman, E. (Eds.), *The Game Design Reader: A Rules of Play Anthology*. 438–459. Cambridge, MA, USA: The MIT Press.

Li, F. W. B., Li, L. W. F. & Lau, R. W. H. (2004). Supporting Continuous Consistency in Multiplayer Online Games. In *Proceedings of the ACM International Conference on Multimedia (Multimedia'04)*, 388–391. New York, NY, USA: ACM. doi: 10.1145/1027527.1027619

Lindley, C. A. (2005). Game Space Design Foundations for Trans-Reality Games. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE '05)*, 397–404. New York, NY, USA: ACM. doi: 10.1145/1178477.1178569

Ling, C. & Salvendy, G. (2005). Extension of Heuristic Evaluation Method: A Review and Reappraisal. *Ergonomia. An International Journal of Ergonomics and Human Factors*, *27* (3), 179–197.

Ling, C. & Salvendy, G. (2008). Effect of Evaluators' Cognitive Style on Heuristic Evaluation: Field Dependent and Field Independent Evaluators. *International Journal of Human-Computer Studies*, *67* (4), 382-393. doi: 10.1016/j.ijhcs.2008.11.002

Linner, D., Kirsch, F., Radusch, I. & Steglich, S. (2005). Context-Aware Multimedia Provisioning for Pervasive Games. In *Proceedings of the IEEE International Symposium on Multimedia (ISM'05)*, 60-68. Washington, DC, USA: IEEE. doi: 10.1109/ISM.2005.46

Lonthoff, J. & Ortner, E. (2007). Mobile Location-Based Gaming as Driver for Location-Based Services (LBS) – Exemplified by Mobile Hunters. *Informatica - An International Journal of Computing and Informatics*, *31* (2), 183–190.

Mack, R. & Nielsen, J. (1993). Usability Inspection Methods: Report on a Workshop Held at CHI'92, Monterey, CA, May 3-4, 1992. *SIGCHI Bulletin*, *25* (1), 28–33. doi: 10.1145/157203.157207

Malone, T. W. (1980). What Makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games. In *Proceedings of the ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems*, 162–169. New York, NY, USA: ACM. doi: 10.1145/800088.802839

Malone, T. W. (1982). Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'82)*, 63–68. New York, NY, USA: ACM. doi: 10.1145/800049.801756

Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S. & Ames, M. (2003). Heuristic Evaluation of Ambient Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*, 169–176. New York, NY, USA: ACM. doi: 10.1145/642611.642642

Markopoulos, P., Read, J., MacFarlane, S. & Höysniemi, J. (2008). *Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers*. San Francisco, CA, USA: Morgan Kaufmann.

McGonigal, J. (2011). *Reality is Broken: Why Games Make Us Better and How They Can Change the World*. New York, USA: The Penguing Books.

McKay, F. & Kölling, M. (2012). Evaluation of Subject-Specific Heuristics for Initial Learning Environments: A Pilot Study. In *Proceedings of the 24th Annual Conference on Psychology of Programming Interest Group (PPIG 2012)*, Paper No. 11. London, UK: London Metropolitan University.

Mohamed, H. & Jaafar, A. (2012). Analyzing Critical Usability Problems in Educational Computer Game (UsaECG). In *Proceedings of the IASTED International Conference on Human-Computer Interaction (HCI 2012)*, 162-168. Calgary, AB, Canada: ACTA Press. doi: 10.2316/P.2012.772-038

Molich, R. & Dumas, J. S. (2008). Comparative Usability Evaluation (CUE-4). *Behaviour & Information Technology*, *27* (3), 263–281. doi: 10.1080/01449290600959062

Molich, R., Kaasgaard, K. & Karyukin, B. (2004). Comparative Usability Evaluation. *Behaviour & Information Technology*, *23* (1), 65–74. doi: 10.1080/0144929032000173951

Molich, R. & Nielsen, J. (1990). Improving a Human-Computer Dialogue. *Communications of the ACM*, *33* (3), 338–348. doi: 10.1145/77481.77486

Montola, M., Stenros, J. & Waern, A. (2009). *Pervasive Games Theory and Design: Experiences on the Boundry Between Life and Play*. Burlington, MA, USA: Morgan Kaufmann.

Moscovich, T. (2007). *Principles and Applications of Multi-touch Interfaces*. Ph. D. Thesis in Department of Computer Science, Brown University, USA.

Muller, M. J., Matheson, L., Page, C. & Gallup, R. (1998). Methods & Tools: Participatory Heuristic Evaluation. *interactions*, *5* (5), 13–18. doi: 10.1145/285213.285219

Mulligan, J. & Patrovsky, B. (2003). *Developing Online Games: An Insider's Guide*. Indianapolis, IN, USA: New Riders.

Munoz, R., Barcelos, T. & Chalegre, V. (2011). Defining and Validating Virtual Worlds Usability Heuristics. In *Proceedings of the 30th International Conference of the Chilean Computer Science Society (SCCC)*, 171-178. Conference Publishing Services. doi: 10.1109/sccc.2011.23

Mäyrä, F. (2008). *An Introduction to Game Studies - Games in Culture*. London, UK: Sage Publications.

Mäyrä, F. (2015). Finland. In Wolf, M. J. P. (Ed.), *Video Games Around the World*. 159-173. Cambridge, MA, USA: MIT Press.

Nardi, B. & Harris, J. (2006). Strangers and Friends: Collaborative Play in World of warcraft. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)*, 149–158. New York, NY, USA: ACM. doi: 10.1145/1180875.1180898

Niedenthal, S. (2009). What We Talk About When We Talk About Game Aesthetics. In *Proceedings of the Digital Games Research Association (DiGRA) 2009 Conference - Breaking New Ground (DIGRA 2009)*, London, UK: Brunel University.

Nielsen, J. (1990). Paper Versus Computer Implementations as Mockup Scenarios for Heuristic Evaluation. In *Proceedings of the IFIP TC13 Third Interational Conference on Human-Computer Interaction (INTERACT)*, 315–320. Amsterdam, The Netherlands: North-Holland Publishing Co.,

Nielsen, J. (1992a). Finding Usability Problems through Heuristic Evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'92)*, 373–380. New York, NY, USA: ACM. doi: 10.1145/142750.142834

Nielsen, J. (1992b). Reliability of Severity Estimates for Usability Problems Found by Heuristic Evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*, 129–130. New York, NY, USA: ACM. doi: 10.1145/1125021.1125117

Nielsen, J. (1993). *Usability Engineering*. London, UK: Academic Press.

Nielsen, J. (1994a). Enhancing the Explanatory Power of Usability Heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*, 152–158. New York, NY, USA: ACM. doi: 10.1145/191666.191729

Nielsen, J. (1994b). Heuristic Evaluation. Chapter 2 In Nielsen, J. & Mack, R. (Eds.), *Usability Inspection Methods*. 25–62. New York, NY, USA: John Wiley & Sons.

Nielsen, J. & Molich, R. (1990). Heuristic Evaluation of User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*, 249-256. New York, NY, USA: ACM. doi: 10.1145/97243.97281

Nokia (2003). *Overview of Multiplayer Mobile Game Design*. Forum Nokia.

Nokia (2004a). *Nokia Series 60 J2ME™ Game Usability Guidelines and Implementation Model*. Forum Nokia.

Nokia (2004b). *Series 40 Developer Platform 1.0: Usability Guidelines For J2ME™ Games*. Forum Nokia.

Nokia (2004c). *Series 60 Developer Platform 2.0: Usability Guidelines For Symbian C++ Games*. Forum Nokia.

Nokia (2005a). *At The Core Of Mobile Game Usability: The Pause Menu*. Forum Nokia.

Nokia (2005b). *Snake Creator Receives Special Recognition From Mobile Entertainment Forum*. http://company.nokia.com/en/news/press-releases/2005/06/16/snake-creator-receives-special-recognition-from-mobile-entertainment-forum (Accessed June, 16, 2005)

Nokia (2006). *Mobile Game Playability Heuristics*. Forum Nokia.

Norman, D. A. (2005). *Emotional Design: Why Do We Love (or Hate) Everyday Things*. New York, NY, USA: Basic Books.

Norman, D. A. (2009). Memory is more important than actuality. *interactions*, *16* (2), 24–26. doi: 10.1145/1487632.1487638

Oulasvirta, A., Tamminen, S., Roto, V. & Kuorelahti, J. (2005). Interaction in 4-Second Bursts: the Fragmented Nature of Attentional Resources in Mobile HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*, 919-928. New York, NY, USA: ACM. doi: 10.1145/1054972.1055101

Overbeeke, K., Djajadiningrat, T., Hummels, C., Wensveen, S. & Frens, J. (2003). Let's Make Things Engaging. Chapter 1 In Blythe, M., Monk, A., Overbeeke, K. & Wright, P. (Eds.), *Funology: From Usability to Enjoyment*. 7–17. Dordrecht, The Netherlands: Kluwer Academic Publisher.

Paavilainen, J. (2010). Critical Review on Video Game Evaluation Heuristics: Social Games Perspective. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology (FuturePlay'10)*, 56–65. New York, NY, USA: ACM. doi: 10.1145/1920778.1920787

Paavilainen, J., Alha, K. & Korhonen, H. (2012). Exploring Playability of Social Network Games. In *Proceedings of the 9th International*

Conference on Advances in Computer Entertainment (ACE'12), 336–351. New York, NY, USA: ACM. doi: 10.1007/978-3-642-34292-9_24

Paavilainen, J., Alha, K. & Korhonen, H. (2015). Domain-Specific Playability Problems in Social Network Games. *International Journal of Arts and Technology*, *8* (4), 282-306.

Paavilainen, J., Korhonen, H., Saarenpää, H. & Holopainen, J. (2009). Player Perception of Context Information Utilization in Pervasive Mobile Games. In *Proceedings of the DIGRA 2009 Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory (DIGRA 2009)*, DIGRA.

Paavilainen, J., Korhonen, H. J. & Alha, K. (2014). Common Playability Problems in Social Network Games. In *Extended Abstracts on 32nd annual ACM conference on Human Factors in Computing Systems (CHI EA'14, Toronto, Ontario, Canada)*, 1519-1524. New York, NY, USA: ACM Press. doi: 10.1145/2559206.2581336

Paddison, C. & Englefield, P. (2004). Applying Heuristics to Accessibility Inspections. *Interacting with Computers*, *16* (3), 507-521. doi: 10.1016/j.intcom.2004.04.007

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L. & Fuller, T. (2003). User-Centered Design in Games. Chapter 53 In Jacko, J. & Sears, A. (Eds.), *Handbook for Human-Computer Interaction in Interactive Systems.* 883–906. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, Inc.

Palm, T. & Koivisto, E. M. I. (2004). Developing Online Mobile Games. In *Proceedings of the Game Developers Conference*, Game Developers Conference.

Pantel, L. & Wolf, L. C. (2002). On the Impact of Delay on Real-Time Multiplayer Games. In *Proceedings of the International workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, 23–29. New York, NY, USA: ACM. doi: 10.1145/507670.507674

Papaloukas, S., Patriarcheas, K. & Xenos, M. (2009). Usability Assessment Heuristics in New Genre Videogames. In *Proceedings of the 13th Panhellenic Conference on Informatics (PCI'09)*, 202–206. Washington, DC, USA: IEEE. doi: 10.1109/PCI.2009.14

Pardo, R. (2006). A Key Note Speech on the Design Philosophy of World of Warcraft. In *Proceedings of the Austin Game Conference*,

Park, N., Lee, K. M., Jin, S.-A. A. & Kang, S. (2010). Effects of Pre-Game Stories on Feelings of Presence and Evaluation of Computer Games. *International Journal of Human-Computer Studies*, *68* (11), 822–833. doi: 10.1016/j.ijhcs.2010.07.002

Peitz, J., Saarenpää, H. & Björk, S. (2007). Insectopia: Exploring Pervasive Games Through Technology Already Pervasively Available. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE '07)*, 107–114. New York, NY, USA: ACM. doi: 10.1145/1255047.1255069

Petrie, H. & Power, C. (2012). What Do Users Really Care About?: A Comparison of Usability Problems Found by Users and Experts on Highly Interactive Websites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'12)*, 2107-2116. New York, NY, USA: ACM. doi: 10.1145/2207676.2208363

Pinelle, D., Wong, N. & Stach, T. (2008a). Heuristic Evaluation for Games: Usability Principles for Video Game Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 1453–1462. New York, NY, USA: ACM. doi: 10.1145/1357054.1357282

Pinelle, D., Wong, N. & Stach, T. (2008b). Using Genres to Customize Usability Evaluations of Video Games. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share (Futureplay '08)*, 129–136. New York, NY, USA: ACM. doi: 10.1145/1496984.1497006

Pinelle, D., Wong, N., Stach, T. & Gutwin, C. (2009). Usability Heuristics for Networked Multiplayer Games. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work (Group '09)*, 169–178. New York, NY, USA: ACM. doi: 10.1145/1531674.1531700

Po, S., Howard, S., Vetere, F. & Skov, M. B. (2004). *Heuristic Evaluation and Mobile Usability: Bridging the Realism Gap*. In Brewster, S. & Dunlop, M. D. (Eds.), International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI), LNCS 3160, 49–60. Berlin Heidelberg, Germany: Springer. doi: 10.1007/978-3-540-28637-0_5

Poels, K., de Kort, Y. & Ijsselsteijn, W. (2007). "It is always a lot of fun!": Exploring Dimensions of Digital Game Experience Using Focus Group Methodology. In *Proceedings of the Conference on Future Play (Future Play '07)*, 83–89. New York, NY, USA: ACM. doi: 10.1145/1328202.1328218

Preece, J., Rogers, Y. & Sharp, H. (2007). *Interaction Design: Beyond Human-Computer Interaction*. New York, NY, USA: John Wiley & Sons.

Rashid, O., Bamford, W., Coulton, P., Edwards, R. & Scheible, J. (2006). PAC-LAN: Mixed-Reality Gaming with RFID-Enabled Mobile Phones. *ACM Computers in Entertainment*, *4* (4), Article No. 4. doi: 10.1145/1178418.1178425

Reeve, J. (2003). *Understanding Motivation and Emotion*. New York, NY, USA: John Wiley & Sons.

Rollings, A. & Adams, E. (2003). *on Game Design*. Indianapolis, IN, USA: New Riders Publishing.

Rouse, R. (2001). *Game Design: Theory and Practice*. Sudbury, MA, USA: Wordware Publishing.

Rubin, J. (1994). *Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests*. New York, NY, USA: John Wiley & Sons.

Röcker, C. & Haar, M. (2006). Exploring the Usability of Video Game Heuristics for Pervasive Game Development in Smart Home Environments. In *Proceedings of the Third International Workshop on Pervasive Gaming Applications (PerGames'06)*, 124–131. http://iis.ipsi.fraunhofer.de/ipsi/ambiente/Liste.asp?AutorID=170

Saarenpää, H., Korhonen, H. & Paavilainen, J. (2009). Asynchronous Gameplay in Pervasive Multiplayer Mobile Games. In *Extended Abstracts on SIGCHI Conference on Human Factors in Computing Systems (CHI EA'09, Boston, MA, USA)*, 4213–4218. New York, NY, USA: ACM Press. doi: 10.1145/1520340.1520642

Salen, K. & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, MA, USA: MIT Press.

Salovaara, A., Johnson, M., Toiskallio, K., Tiitta, S. & Turpeinen, M. (2005). Playmakers in Multiplayer Game Communities: Their Importance and Motivations for Participation. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE '05)*, 334–337. New York, NY, USA: ACM. doi: 10.1145/1178477.1178540

Sanchez-Crespo, D. (1999). *Learn Faster to Play Better: How to Shorten the Learning Cycle*. http://www.gamasutra.com/features/19991108/dalmau_01.htm (Accessed November 5, 2013)

Schaffer, N. (2007). *Heuristics for Usability in Games*. White Paper. Available at: http://gamesqa.files.wordpress.com/2008/03/heuristics_noahschafferwhitepaper.pdf (Accessed August 14, 2007)

Schell, J. (2008). *The Art of Game Design*. Burlington, MA, USA: Morgan Kaufmann.

Sears, A. (1997). Heuristic Walkthroughs: Finding the Problems Without the Noise. *International Journal of Human-Computer Interaction*, *9* (3), 213–234. doi: 10.1207/s15327590ijhc0903_2

Seay, A. F., Jerome, W. J., Lee, K. S. & Kraut, R. E. (2004). Project Massive: A Study of Online Gaming Communities. In *Extended Abstracts on SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA'04, Vienna, Austria)*, 1421–1424. New York, NY, USA: ACM Press. doi: 10.1145/985921.986080

Shelley, B. (2001). *Guidelines for Developing Successful Games*. http://www.gamasutra.com/features/20010815/shelley_01.htm (Accessed November 5, 2013)

Shneiderman, B. & Plaisant, C. (2009). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Boston, MA, USA: Addison-Wesley.

Sim, G., Read, J. & Cockton, G. (2009). *Evidence Based Design of Heuristics for Computer Assisted Assessment*. In Gross, T., Gulliksen, J., Kotze, P., Oestreicher, L., Palanque, P., Oliveira Prates, R. & Winckler, M. (Eds.), IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT), LNCS 5726, 204-216. Berlin Heidelberg, Germany: Springer. doi: 10.1007/978-3-642-03655-2_25

Sim, G., Read, J. & Holifield, P. (2008). Heuristics for Evaluating the Usability of CAA Applications. In *Proceedings of the 12th CAA International Computer Assisted Assessment Conference*, 283-294. Loughborough University.

Smith, H. (2003). *Orthogonal Unit Differentiation*. http://www.docstoc.com/docs/42438381/Orthogonal-Unit-Design (Accessed May 13, 2014)

Snow, B. (2007). *Gaming Usability 101*. http://www.businessweek.com/print/innovate/content/oct2007/id20071012_041625.htm (Accessed December 7, 2008)

Somervell, J. P. & McCrickard, D. S. (2004). Comparing Generic vs. Specific Heuristics: Illustrating a New UEM Comparison technique. In

Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting (HFES '04), 2480–2484. Santa Monica, CA, USA: Human Factors and Ergonomics Society. http://people.cs.vt.edu/~mccricks/papers/sm-hfes04.pdf

Sorens, N. (2007). *Rethinking the MMO*. http://www.gamasutra.com/view/feature/1583/rethinking_the_mmo.php?page=3 (Accessed February, 25, 2011)

Sotamaa, O. (2002). All The World's A Botfighter Stage: Notes on Location-based Multi-User Gaming. In *Proceedings of the Computer Games and Digital Cultures Conference (CGDC)*, 35–44. Tampere, Finland: Tampere University Press.

Studio 7.5 (2005). *Designing for Small Screens*. Lausanne, Switzerland: AVA Publishing SA.

Suomela, R. & Koivisto, A. (2006). *My Photos Are My Bullets - Using Camera as the Primary Means of Player-to-Player Interaction in a Mobile Multiplayer Game*. In Harper, R., Rautenberg, M. & Combetto, M. (Eds.), International Conference on Entertainment Computing (ICEC), LNCS 4161, 250–261. Berlin Heidelberg: Springer. doi: 10.1007/11872320_30

Sutcliffe, A. & Gault, B. (2004). Heuristic Evaluation of Virtual Reality Applications. *Interacting with Computers*, *16* (4), 831-849. doi: 10.1016/j.intcom.2004.05.001

Sweetser, P. & Wyeth, P. (2005). GameFlow: A Model for Evaluating Player Enjoyment in Games. *ACM Computers in Entertainment*, *3* (3), Article 3A. doi: 10.1145/1077246.1077253

Tamminen, S., Oulasvirta, A., Toiskallio, K. & Kankainen, A. (2004). Understanding Mobile Context. *Personal and Ubiquitous Computing*, *8* (2), 135–143. doi: 10.1007/s00779-004-0263-1

Tan, E. S. & Jansz, J. (2008). The game experience. Chapter 23 In Schifferstein, H. N. J. & Hekkert, P. (Eds.), *Product Experience*. 531–556. San Diego, CA, USA: Elsevier.

Taylor, T. L. (2006). *Play Between Worlds: Exploring Online Game Culture*. Cambridge, MA, USA: MIT Press.

Theng, Y.-L., Ho, A. & Wee, E. (2008). *Exploring Factors That Make Online Interactive Games Successful: A Heuristic Approach*. In Lee, S., Choo, H., Ha, S. & Shin, I. C. (Eds.), Asia-Pacific Conference on Computer-Human Interaction (APCHI), LNCS 5068, 362–371. Berlin Heidelberg: Springer. doi: 10.1007/978-3-540-70585-7_41

Thomas, S., Schott, G. & Kambouri, M. (2003). *Designing for Learning or Designing for Fun? Setting Usability Guidelines for Mobile Educational Games*. In Attewell, J., Savill-Smith, C. & Attawell, S. (Eds.), Learning with Mobile Devices, (MLEARN 2003), 173–181. London, UK: Learning and Skills Development Agency.

Usability_Glossary (2002). *Playability*. http://www.usabilityfirst.com/glossary/playability/ (Accessed June 8, 2005)

Warner, D. E. & Raiter, M. (2005). Social Context in Massively-Multiplayer Online Games (MMOGs): Ethical Questions in Shared Space. *International Review of Information Ethics*, *4*, 46–52. http://www.i-r-i-e.net/inhalt/004/Warner-Raiter.pdf

Vermeeren, A. P. O. S., van Kesteren, I. E. H. & Bekker, M. M. (2003). Managing the Evaluator Effect in User Testing. In *Proceedings of the IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT)*, 647–654. Amsterdam, The Netherlands: IOS Press.

Williams, P., Nesbitt, K. V., Eidels, A. & Elliott, D. (2011). Balancing Risk and Reward to Develop an Optimal Hot-Hand Game. *The International Journal of Computer Game Research*, *11* (1), http://gamestudies.org/1101/articles/williams_nesbitt_eidels_elliott

Witkin, A. & Baraff, D. (1997). *Physically Based Modeling: Principles of Practice (Online Siggraph'97 course notes)*. http://www.cs.cmu.edu/~baraff/sigcourse/ (Accessed April 5, 2013)

Xu, C.-w. (2008). A Software Framework for Online Mobile Games. In *Proceedings of the International Conference on Computer Science and Software Engineering (CSSE)*, 558–561. Washington, DC, USA: IEEE. doi: 10.1109/csse.2008.1412

Yee, N. (2006). Motivations for Play in Online Games. *CyberPsychology & Behavior*, *9* (6), 772-775. doi: 10.1089/cpb.2006.9.772

Zagal, J. P., Ladd, A. & Johnson, T. (2009). Characterizing and Understanding Game Reviews. In *Proceedings of the 4th International Conference on Foundations of Digital Games (FDG '09)*, 215-222. New York, NY, USA: ACM Press. doi: 10.1145/1536513.1536553

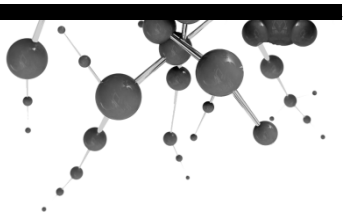
Zagal, J. P., Nussbaum, M. & Rosas, R. (2000). A Model to Support the Design of Multiplayer Games. *Presence*, *9* (5), 448–462. doi: 10.1162/105474600566943

Zazelenchuk, T. (2006). *Heuristic Evaluation and Its Alternatives*. http://www.userfocus.co.uk/articles/heuristics.html (Accessed January 8, 2013)

Zhang, J., Johnson, T. R., Patel, V. L., Paige, D. L. & Kubose, T. (2003). Using Usability Heuristics to Evaluate Patient Safety of Medical Devices. *Journal of Biomedical Informatics*, *36* (1-2), 23-30. doi: 10.1016/S1532-0464(03)00060-1

Zhang, Z., Basili, V. & Shneiderman, B. (1999). Perspective-Based Usability Inspection: An Empirical Validation of Efficacy. *Empirical Software Engineering*, *4* (1), 43–69. doi: 10.1023/A:1009803214692

Zuk, T., Schlesier, L., Neumann, P., Hancock, M. S. & Carpendale, S. (2006). Heuristics for Information Visualization Evaluation. In *Proceedings of the AVI workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV '06)*, 1–6. New York, NY, USA: ACM. doi: 10.1145/1168149.1168162



Paper I

Korhonen, H. & Koivisto, E. M. I. (2006). Playability Heuristics for Mobile Games. In *Proceedings of International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'06, Helsinki, Finland)*, 9-16. New York, NY, USA: ACM. doi: 10.1145/1152215.1152218

Playability Heuristics for Mobile Games

Hannu Korhonen
Nokia Research Center
Visiokatu 1
33720 Tampere, Finland
hannu.j.korhonen@nokia.com

Elina M.I. Koivisto
Nokia Research Center
Visiokatu 1
33720 Tampere, Finland
elina.m.koivisto@nokia.com

ABSTRACT
Expert evaluation is a widely used method for evaluating the usability of software products. When evaluating games, traditional usability heuristics lack comprehension and cannot be directly applied. In this paper, we introduce playability heuristics that are specifically designed for evaluating mobile games. Heuristics form a core model that can be used in any mobile game evaluation. The model consists of three modules: *Game Usability, Mobility,* and *Gameplay*. The mobile context has some unique characteristics, which require special attention during the evaluation. These characteristics are described in mobility heuristics. Mobile devices also set some of their own requirements for general usability and these issues are described along with game usability heuristics. These heuristics have been developed by using an iterative design process of a mobile game. In addition, we have validated the heuristics and evaluated five mobile games by using them with the expert evaluation method. The results indicate that playability problems, which violate game usability or mobility heuristics, are quite easy to identify. Gameplay problems are harder to find, but gameplay heuristics help in evaluation and focus on different aspects of the gameplay.

Categories and Subject Descriptors
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces

General Terms
Design, Human Factors

Keywords
Mobile Games, Playability, Heuristics, Evaluation, Usability.

1. INTRODUCTION
The usability of a software product is often defined as effectiveness, efficiency, and user satisfaction in a specified context of use [7]. Games, however, are most enjoyable and fun when they provide sufficient challenge for a player. The challenge can be, for instance, in learning the game, solving problems or discovering new things.

A good gaming experience requires a lot from the user interface. It should be convenient, reliable, and usable so that the player can concentrate on playing the game and enjoying it instead of struggling with the user interface. In addition, the game design itself has a huge impact on the gaming experience. If the rules or game world contains implausible features, the players can be easily offended or frustrated and quit playing the game.

Playing games with mobile devices in a mobile context is quite a new research area and the heuristics for evaluating these aspects has not been discussed in previous playability heuristics. We have developed playability heuristics to cover the mobile context aspects. The playability heuristics can be used with an expert evaluation method to identify possible playability problems in the user interface and game design in early phase of a game project. These heuristics cover general usability, mobility, and gameplay issues of the game. This method is mainly targeted at pre-production and production phases of a game project, but it can also be utilized in post-production phase.

Mobile games (see for example, Figure 1) have become more and more complex over the years and currently they resemble console and PC games. We have used these heuristics to evaluate different game titles that are designed for smartphones[1] and mobile gaming devices, such as Nokia N-Gage[2]. These heuristics, however, are suitable for evaluating games in other platforms as well since gameplay and game usability are common for all games.

Figure 1 An example of a mobile game on smartphones

MobileHCI'06, September 12–15, 2006, Helsinki, Finland.

[1] Smartphones combine telephone and PDA functions, and enable the user to install additional applications on the device. In addition, the device uses particular operating system such as Symbian, Windows Mobile or Linux.

[2] http://web.n-gage.com/en-US/gamedeck/ngage_qd/

2. RELATED WORK

Usability heuristics are designed for evaluating the user interface of the application. Users typically have their own goals that they try to accomplish with the application and the evaluations try to find out how easily and efficiently the user can achieve these goals. Probably the most commonly used usability heuristics are those originally developed by Nielsen and Molich [15]. Nielsen has then refined these heuristics [13], resulting in the current list of usability heuristics [14]. Additions have also been made to Nielsen's heuristics when evaluating utility software. For example, Muller *et al.* have added three new heuristics that they have considered to be useful [11].

Games differ from utility software in some key characteristics. In games, the purpose is to have fun and enjoy playing the game. Learning to play the game, solving problems, or discovering new things is part of that experience. Moreover, in a game, the players do not know in advance what to expect. Game designers have created the game content and defined goals that the players must achieve. Playing a game is not straightforward either, but it is challenging, and the player needs to work towards goals. Therefore, applying general usability heuristics in game evaluations is not sufficient and using only them would leave many important aspects of the game unprocessed [5].

There are other playability heuristics already available. Malone created the first heuristics for evaluating educational games [10]. More recently, Federoff has created a list of heuristics as a result of a case study at a game development company [5]. Desurvire *et al* have created heuristics that are best suited for evaluating general issues in early development phase with a prototype or mock-up [3]. Järvinen *et al* have developed a theoretical tool for evaluating playability of games trough studying such notions and concepts as 'optimal experience', 'playability', and 'gameplay' [9]. We considered using these heuristics or tools in our evaluations, but they were not feasible. First, the existing heuristics did not deal with mobility issues, which is one of our main targets. Second, all heuristics were not described in detail so that they could have been directly adapted to our process. Third, some of the heuristics were overlapping, which made them ambiguous. Therefore, we decided to start developing our own set of heuristics, which would overcome these shortcomings.

3. DEVELOPMENT OF PLAYABILITY HEURISTICS FOR MOBILE GAMES

We started the development of the playability heuristics for mobile games by defining what kinds of aspects should be evaluated. Even if our original purpose was to concentrate on the gameplay issues, we quickly noticed that game usability is so closely related to the gameplay that in game evaluation general usability aspects cannot be ignored. General usability of the mobile games is a very important aspect and players do not want to struggle with it, because players are not interested in the user interface. The user interface of the game and control keys that are used for controlling the game characters should be very natural and intuitive to use. Since we are evaluating mobile games, another aspect is to evaluate how well a game fits into the mobile context. The third aspect is to evaluate the actual content of the game. This is also the characteristic that differentiates games from other software. Unlike utility software, game developers have created game content and all tasks that players need to achieve in the game. Normally with utility software, users define they own goals and content that they will manage. The software is a tool and its responsibility is to help the users to achieve their goals as fast and efficiently as possible.

Our evaluation model for mobile games is modular and consists of three core modules: *Gameplay*, *Mobility*, and *Game Usability* (Figure 2). They are common for any mobile game. In addition, *Game Usability* and *Gameplay* modules are generic and they can be used for evaluating any mobile game regardless of the platform. The number of modules is smaller than what Federoff and Desurvire *et al* have proposed [5],[3]. The reason for doing this is that we have consolidated some of their modules. For example, the *Gameplay* module incorporates game mechanics because they belong together inextricably. Depending on the point of view many heuristics in these two groups can belong to either group. Gameplay tells about the structures of the player interaction with the game system and with other players in the game [1]. Game mechanics, on the other hand, consists of rules that define the operation of the game world and make up the core mechanics, the foundations of gameplay [16]. Gameplay occurs when the player interacts with the game mechanics and possibly other players. In our model the *Gameplay* module contains also game story, which is separate in Desurvire's list.

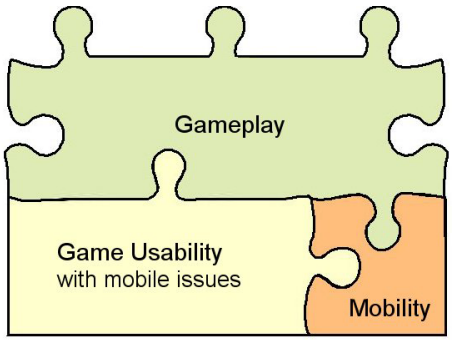


Figure 2 Modules in the core playability model

The structure of the model defines relations of these modules. Gameplay is "the heart of the game" and in order to evaluate it properly there should not be any major playability problems in game usability nor mobility. Game usability and mobility have also correlation in mobile games. Even though *Mobility* is a separate module in the model, there are also mobile issues in game usability heuristics that need to be taken into account when evaluating game usability in mobile games.

A modular structure suggests that it is possible to use each of these modules separately and evaluate the game against one module at a time. This is very useful when evaluating early versions of the game. Gameplay should be evaluated already in an early design phase when there are design documents available. Game usability and mobility will increase their importance when there is interaction design or prototypes available. The structure also suggests that it is possible to add new modules if needed. Especially the gameplay module may need some supplements with different game styles because current gameplay heuristics are very general and applicable for any game. This also helps maintaining the number of heuristics reasonable.

3.1 Defining Mobility Heuristics

We started the development of the mobility heuristics by analyzing mobile phones and their context of use. We analyzed how the context, in which the mobile phones are used, affects the tasks that the user does and in what kind of context the mobile phones are normally used. Since mobile games are also mobile applications, it is presumable that similar requirements would also apply for them. In addition, physical characteristics of the mobile phones probably have some influence to mobile gaming as well.

The mobile context can be very different from the office context. Users will use their mobile phones outdoors where lighting conditions and noise can change frequently. Occasionally users may need to observe their surroundings while doing something with their phones. In addition, when using a mobile phone in public places, there can be other persons in the vicinity, who must be taken into consideration.

A mobile phone is an excellent companion for killing time or just doing something during short breaks because it is always with the user. Taking a photo, sending a message, checking the calendar, or browsing a web site are typical tasks that should be initiated without delay. Therefore, the application and the phone should be in operating mode instantly.

A mobile phone is a multi-purpose device, which is, however, used first and foremost for communication. Incoming calls or messages will inevitably cause interruptions for other use of the device. From our experience phone calls are still in high priority and users will answer an incoming call. Messages are not ranked as high and reading a received message can wait until the other task is completed, unless the user is expecting the message.

In the mobile context, interruptions can be triggered by an external event. These events require the user's attention and the task that the user was doing with the mobile phone must be suspended. Examples of such interruptions may be an encounter of a friend or arrival of the bus. The device itself has no means to prepare or act on such situations, but it can let the user decide what to do.

Characteristics of mobile devices will also set some requirements for mobile applications. Users interact with applications by using a standard 12-key keypad and few navigation keys. In some cases they can use a miniature-size joystick, which is operated with the thumb. The small screen size, insufficient audio capabilities, limited processing power, and battery limitations will cause additional requirements that need to be taken into account when designing mobile applications.

3.2 Initial Playability Heuristics

Next we will present heuristics derived from the analysis and discuss the observations. Results of the mobile context analysis, a review of Nielsen's heuristics and game design guidelines helped us create the first version of playability heuristics for mobile games. This version contained the following heuristics:

H1: Don't waste the player's time

H2: Prepare for interruptions

H3: Take other persons into account

H4: Follow standard conventions

H5: Provide gameplay help

H6: Differentiation between device UI and the game UI should be evident

H7: Use terms that are familiar to the player

H8: Status of the characters and the game should be clearly visible

H9: The Player should have clear goals

H10: Support a wide range of players and playing styles

H11: Don't encourage repetitive and boring tasks

We tried to keep the number of heuristics minimal because a smaller number would be easier to use and remember during evaluation. The other objective was to get the heuristics generic enough so they would cover as many problems as possible.

Heuristics H1, H2, and H3 are derived from the analysis of mobile phone and mobile context. *"Don't waste the player's time*" comes from the observation that mobile phones are used for short periods for killing time or doing something useful and there is not much time for waiting. "*Prepare for interruptions*" heuristic combines characteristics of the multi-purpose device and its use in the mobile context. This heuristic could actually be used when evaluating any mobile application because interruptions can happen when the user is using non-game applications as well. "*Take other persons into account*" applies for all mobile devices because they are used in public places and other people are almost always in the vicinity. The player should not end up in an embarrassing situation where disturbance is inappropriate.

Although mobile phones are quite new devices as a gaming platform, certain conventions can already be drawn from current games. For instance, number five on the mobile phone's keypad works in many cases as a selection key. The device's standard input methods should be used for controlling the game [6]. Another interesting point is the form factor of the device. If the device is meant to be operated one-handed, the game should be playable with one hand. "*Follow standard conventions*" states that a game should follow these conventions because it makes using the device and learning the game controls easier.

Heuristics H5, H7, and H8 are modified versions of general principles for user interface design that Nielsen and Molich have developed [14]. "*Provide gameplay help*" is based on the fact that players rarely read manuals before they start playing the game and carrying a paper manual around would be very inconvenient. "*Use terms that are familiar to the player*" means that the game should avoid terminology that is unfamiliar to a player. The player may not be aware or even interested in knowing differences of network connections nor their associated settings. Also, if the game uses some abbreviations or terms, they should be introduced to the player when they are first encountered. "*Status of the characters and the game should be clearly visible*" is another modified principle from general usability heuristics. The player should always know the current state of the game, for example, whose turn it is to make the next move or what is the current condition of the character. Nothing is more frustrating than unexpected game ending without knowing why it happened.

Games rely much on the immersion during a play session. The game world is a place where the player wants to concentrate on

events that are happening over there. Using phone UI widgets to display some information will break this immersion. "*Differentiation between device UI and the game UI should be evident*" states that game developers should not use device UI, but create a new UI for the game. Preferably the game should use full screen mode and not to use any device's user interface widgets in the game interface in order to maintain the immersion.

Heuristics H9, H10 and H11 are purely gameplay oriented heuristics. "*The Player should have clear goals*" is a fundamental issue in all games. In order to play the game, a player should understand the goals that exist in the game [4]. According to the Flow theory, having a clear goal in mind is a core of an enjoyable experience [2]. Within a non-gaming context, users create their own goals and use applications to achieve these goals. When playing games, however, a player does not know in advance what to expect. A game designer has created the game content and defined goals that the player must achieve. Even though this heuristic does not mean that the game must tell how to achieve goals, it should tell what the goal is and how to begin. In some games players can also create their own goals or select a goal from the pre-defined list of goals. "*Support a wide range of players and playing styles*" means that the game should provide at least a possibility to adjust the game challenge according to the player's experience. The players can vary a lot in terms of both experience and preferred play styles. Usually this can be done with difficulty levels that change gameplay. "*Don't encourage repetitive and boring tasks*" deals with the issues that sometimes make games boring or frustrating. The game should not require repetition of tasks without changing any conditions. Repeating the same tasks over and over again is often called tread milling or grinding, and is usually a guaranteed way of killing the fun in the game. However, it should be noted that a training phase in the game is not grinding because the player needs to practice basic actions, for instance, how the character is controlled in the game. During the training phase, it is useful to repeat certain tasks so that the player learns them.

3.3 Expert Evaluation of Games

In the expert evaluation of utility software, 1-6 evaluators explore the application and write a report about findings that violate heuristics. In addition, the report should contain design solutions that are successful in the application. This should prevent designers to change features, which are working well. Usually the evaluators are usability experts, but it is recommended that they have some kind of domain expertise as well because it helps with the evaluation [14]. In game evaluations, the domain expertise is mandatory. At least the evaluators should be interested in games and preferably be familiar with the game genre of the tested game. The reason for requiring knowledge about the game genre is well grounded because each genre has its own characteristics and conventions that should be followed in order to create a successful game. Also some our gameplay heuristics are not relevant for all game styles.

An expert evaluation of the application can take a couple of hours at minimum. The evaluators go through the application a couple of times and write a report about their findings. When we have conducted evaluations with mobile games, we have noticed that the time required to do a game evaluation is much longer. First, the evaluators have to learn to play the game until all of its aspects can be studied. In addition, due to the nature of games as entertaining products, the whole game is not usually revealed to the player in the beginning. Instead, the player discovers new things in the game little by little by playing the game.

Games are often designed so that there is about to 20-40 hours of playing time in a game. This applies mainly to games which have a storyline. For games without a strict storyline, it can be considerably longer. Of course, the evaluators can examine common things, such as the user interface and general usability of the game, in a shorter time, but evaluating the gameplay requires that the evaluators have really played the game.

3.4 Validation of the Heuristics

After formulating this initial set of heuristics, we evaluated a mobile game (Game A_1) using these heuristics as a guideline. The game was in production phase and the first (alpha) version was soon to be finished. The game was playable on smartphones. The evaluation was done by four experts. One evaluator had game design experience and the other evaluators had done normal utility software evaluations. Two of them were eager game players during their spare-time.

During the evaluation, it came apparent that the list is not sufficient, and many playability problems would have left unidentified with only these heuristics. We found 61 playability problems, but for 16 playability problems we did not have a proper heuristic. The list did not cover all user interface and general usability problems, but more specifically was missing gameplay heuristics. In addition, we found four playability problems that were related to the multiplayer features of the game.

The number of playability problems that violated each heuristic is illustrated in Figure 3. The game design succeeded quite well in the mobility aspects and there were only six playability problems that violated the first three heuristics. Unfortunately, all of them were critical problems that required immediate correction.

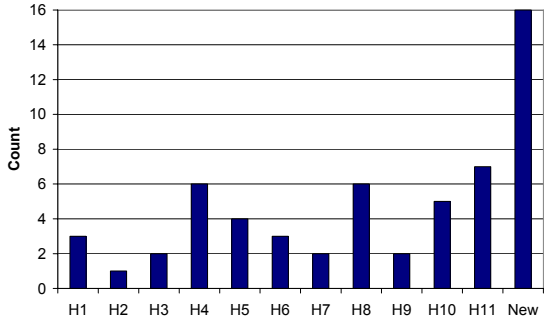


Figure 3 Evaluation results from the first evaluation

The game had many problems in screen layout design and basic navigation. The game UI design tried to follow device UI structure even though there were more commands available than what was possible with device's UI style. This caused confusing moments and players did not know how the control keys worked. In addition, some of the commands were scattered around the screen and the player should have followed certain navigation paths to reach all commands. These playability problems violated heuristics H4 and H6 constantly.

Another severe playability problem in the game was that is was not easy to notice the current status of the game and the player's

characters (Heuristic H8). This was a really severe problem because the pace of the game was fast and the player was forced to make quick decisions for the next move. Furthermore, the player needed to repeat same commands to the characters over and over again, which made the combat system cumbersome.

3.5 Re-evaluating the Heuristics

During the game evaluation we found 16 playability problems for which we did not have a proper playability heuristic. We started to analyze these problems and noticed that most of them were related to gameplay issues, but there were also issues related to the user interface and general usability. Moreover, we had identified four playability problems that were related to multi-player features of the game.

Gameplay is a very complicated part in game evaluation and our initial list did not cover it thoroughly. Although players should have clear goals in the game, they should be able to see progress towards a goal and be rewarded when reaching the goal. In addition, players should be able to compare the achievements. During the first evaluation round we noticed that it is also very important for the player to be in control. It is, after all, the player's responsibility to decide what to do or how to do it, even though game designers have created those events. In addition, if the game contains a story, it should support the gameplay and be meaningful to the player.

All players are novice players in the beginning. During the first play sessions, the player will create a first impression of the game, which is very difficult to change. After that the player gains experience with the game and the game should be able to adjust the challenge for experienced players as well. Balancing the challenge, strategies and pace for various players is a key point for all games. Gaming experience will also create different playing styles, which game designers may not have even thought about.

During the play session it is important that the game world is consistent and every item has purpose. Players are particularly smart in trying different approaches when they are solving problems or discovering new things. Sometimes combining different strategies in the game may result unexpected situations even for game developers[3]. In these situations the game should not stagnate, but the players should be able to continue playing or conclude the play session. Unexpected situations can also lead to exploiting [12].

Game usability issues are also more extensive than we assumed at first. Designing an efficient screen layout, which contains all necessary data, is not an easy task especially on small screens. In addition, due to rapid development of graphics cards, games look visually appealing and players expect that too. However, the game graphics should always support the gameplay and story. Audio is another feature that is left unnoticed in many designs. In mobile games the importance of music and sound effects increases because they create sound environment for the game, but it can also disturb other persons in the vicinity.

In order to provide a satisfying gaming experience, the game should use game controls that are convenient and flexible and provide feedback on player's actions. Moreover, the game should not require the player to remember things unnecessarily or allow the player to make irreversible errors with the user interface. These may seem like irrelevant issues, but during the play session the player does not want to struggle with the user interface, but instead concentrate on playing the game.

As a result, we defined 18 new heuristics to be included in our original heuristics list ending up total of 29 heuristics. Before the next validation round, we arranged a review of the playability heuristics with experienced game designers. Based on their comments we made some modifications to the heuristics.

3.6 Revised Playability Heuristics

Our current list of playability heuristics contains following heuristics categorized into three groups. Detailed descriptions of each heuristic are provided in an article discussing mobile game playability heuristics [8].

3.6.1 Game Usability

The game usability heuristics (Table 1) cover the game controls and interface through which the player interacts with the game. As a general rule, the game interface should allow the player to control the game fluently and display all necessary information about the game status and possible actions. The game interface is usually the first thing that a player encounters when starting to play a new game. Good game usability ensures that the player will have another enjoyable play session.

Game usability heuristics can be grouped to several subgroups. Heuristics GU1-GU5 are related to visual design and how information is presented. This includes also terminology that is used in the game. Heuristics GU6-GU8 will deal with how navigation is arranged and what controls are used for navigation and controlling the game character. The rest of the heuristics are related to other important aspects like getting feedback and how the game can help or guide the player to concentrate on playing the game.

Table 1 Heuristics for evaluating game usability

No.	Game Usability Heuristics
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

[3] This is also called "emergent gameplay".

3.6.2 Mobility

While the game usability heuristics deal with the user interface issues, the mobility heuristics (Table 2) concern issues that affect mobility of the game. Since mobile devices do not dictate where and when games are played, the game design should assimilate this freedom into the game experience. Mobility is defined by how easily the game allows a player to enter to the game world and how it behaves in diverse and unexpected environments.

Table 2 Heuristics for evaluating mobility

No.	Mobility Heuristics
MO1	The game and play sessions can be started quickly
MO2	The game accommodates with the surroundings
MO3	Interruptions are handled reasonably

3.6.3 Gameplay

Gameplay heuristics (Table 3) are valid regardless of the platform on which the game is played. When evaluating gameplay, it is recommended that evaluators have at least some game design expertise. They should also understand the design goals and know the target players. The evaluators need not belong to the target group themselves, but it is imperative to get more familiar with it.

Table 3 Heuristics for evaluating gameplay

No.	Gameplay Heuristics
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 differentiation[4]
GP 14	The player does not lose any hard-won possessions

[4] Units in the game should be designed so that they are functionally different

4. VALIDATING THE HEURISTICS

In order to validate these heuristics we have evaluated several mobile games during last months. Each game was evaluated by two to four evaluators. One evaluator was always a usability expert. Other evaluators were game designers with basic knowledge on usability issues and the evaluation method.

4.1 Games

We selected five games for the evaluations. Game styles and target players were different in order to see how well the heuristics can identify playability problems. Characteristics of games are listed in Table 4. All games were developed by different game companies, and they have not been published yet. Usually we evaluated the first alpha version of the game, which means that most of the features were implemented and the game was running on a mobile device. For one game we evaluated single features because the evaluation happened before the alpha milestone.

Table 4 Game Characteristics

	Game Style	Player mode	Target Player	Device	Evaluators
A_1	Combat	Multi-player	20+, Male	Smart phone	4
A_2	Combat	Multi-player	20+, Male	Mobile Gaming Device	2
B	Adventure	Multi-player	18+, Male	Smart phone	3
C	Simulation	Single Player	12+, Female	Mobile Gaming Device	2
D	Puzzle	Single Player	10+, Neutral	Mobile Gaming Device	3

In game A_1 two players fight duels and collect game characters. Game A_2 is the second version of the game A_1 after corrections based on the first evaluation round. The game felt like a new game after these corrections. In game B a player controls ships and battles on the sea. Game C is a typical simulation game where a player observes and controls game characters indirectly. In game D a player solves puzzles in the game world, which consists of different levels.

4.2 Game Usability and Mobility Heuristics

Game usability issues were the easiest playability problems to be identified in the games. Unfortunately, they seem to be the easiest heuristics to be violated too. At least four games violated each heuristic in the *Game Usability* module. The only exception was preventing a player to make irreversible errors (GU10). Games A_1, B, and C violated this heuristic, but only in a certain situation. During the evaluations we identified 151 playability problems related to game usability (Figure 4). It is noticeable that quite often playability problems in a game were concentrated in certain heuristic. For example game A_1 suffered from navigation problems (GU6), whereas games C and D had many terminology related problems (GU5). Memorizing different things (GU11) in the user interface was a problem in games A_2 and B.

Designing an efficient and visually pleasing user interface is not an easy task in mobile devices. Displaying sufficient amount of data and required commands on the screen is challenging. A game user interface tends to keep everything visible and arranging them reasonably is difficult. Game A_2 in particular suffered these kinds of problems, but the visual design was probably not yet finished.

A small number of mobility problems were expected to be found because the games were designed to be mobile. However, all games, except game C had some playability problems in this area. We identified 10 playability problems related to mobility (Figure 4). These heuristics are quite easy to test since an evaluator needs a Subscriber Identity Module (SIM) card for a device and test how interruptions are handled (MO3). Accommodation to the surroundings (MO2) and launching the game (MO1) are easily tested.

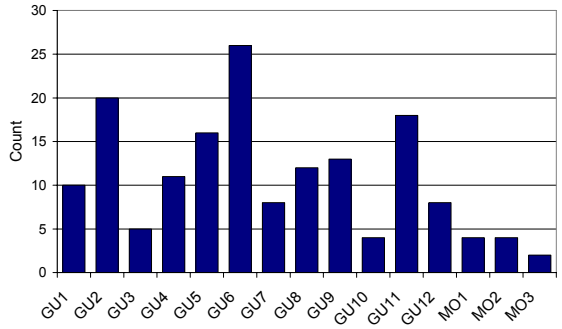


Figure 4 Playability problems violating game usability and mobility heuristics

4.3 Gameplay Heuristics

Gameplay is the most difficult aspect to evaluate because it requires that evaluators explore all aspects in the game. Evaluating gameplay issues may take considerable amount of time. The game complexity will also affect the time needed greatly. We found 64 gameplay problems from the games (Figure 5). The most common violations of gameplay heuristics focused on heuristics GP1, GP3, GP4, and GP5. All evaluated games violated these heuristics. Surprisingly, games had problems in defining understandable goals. This should be the core design concern in game development projects since otherwise players do not know what to do in the game. These kinds of problems in a game may be an indication of design flaws. Other difficult aspects to design were letting the player be in control and balancing challenge, strategy and pace in the game.

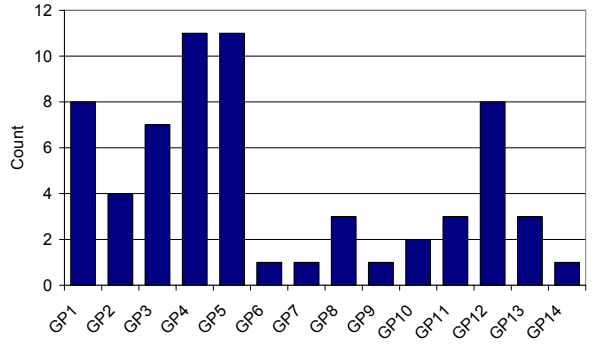


Figure 5 Playability problems violating gameplay heuristics

Another heuristic that was violated by all games except game C was GP12 (The game is consistent). Risk of having these kinds of problems increases when the game becomes more complex.

Gameplay problems related to heuristics GP8, GP9, GP10, and GP11 were found only in Game A_1. The reason for this is that the evaluation time for game A_1 was substantially longer than for other games and the evaluators were able to explore the game thoroughly. Especially, finding these kinds of gameplay problems in the game requires that the evaluators are really familiar with the game and are able to try unique things during the gameplay.

5. DISCUSSION AND FUTURE WORK

Playability heuristics proved to be very efficient in game evaluations and games had playability problems associated to each heuristic. We found 235 playability problems from five mobile games. Some of the problems were more severe than others, but all of them still have some influence on the overall gaming experience. Playability heuristics helped evaluators focus on specific aspects in the game. Even though the evaluators were not able to find that many playability problems from gameplay that violated specific heuristic, it does not necessary imply that the heuristic is useless. On the contrary, this is usually a positive result, because game designers have then succeeded in the design phase and evaluators remembered to acknowledge that aspect as well.

During the evaluation, the most easily identifiable playability problems were related to game usability and mobility. Especially, information visualization and navigational problems were easy to spot. On the other hand, evaluating these aspects from the game resembles normal usability evaluations of utility software. Evaluating the gameplay is much harder. In order to evaluate the gameplay properly requires that game usability issues are checked beforehand. In the future, we should start looking for guidelines that would help finding gameplay problems and thus, making the evaluation more efficient.

Another task for future work is to define playability heuristics for multi-player features. During these evaluations we had three games that are multi-player games. Evaluation revealed some playability problems that do not fit into any existing module. This is the main reason why we designed our heuristics modular. Gameplay heuristics are meant to be general for all game styles and for example multi-player heuristics are applicable to multi-player games only. Therefore, they cannot be included in the gameplay heuristics. However, our model allows other game style specific heuristics to be included in the core heuristics and to be used in game evaluation.

We will continue validating heuristics and evaluate more mobile games using these heuristics. By doing this we will get enough data as to whether these heuristic are applicable for all mobile games and how much there is need for game style specific heuristics. Comparing results from playability testing and expert evaluation would also help in validating these heuristics. In playability testing, players from the target group play the game. Results from these methods should show whether similar problems are found and which method should be used to test different versions of the game.

We will provide these heuristics with descriptions to the public. Any game evaluator or game designer can use them in their evaluations. It is important to know whether these heuristics are

understandable for evaluators that do not have previous experience with them. Another interesting aspect is to see how these heuristics can help during the game design phase because every major change is easier to do when a feature is only written in the design documents.

6. CONCLUSION

We have introduced playability heuristics for mobile games. These heuristics form a core model, which can be used for evaluating any mobile game. The model consists of three modules: *Game Usability, Mobility*, and *Gameplay*. The *Game Usability* module covers the game controls and interface through which the player interacts with the game. Also, it contains common usability aspects that help the player to get into the game and interact with it. *Mobility* module concerns issues that make the game mobile. The *Gameplay* module deals with issues that arise when the player interacts with the game mechanics and story.

We have developed the first version of heuristics in parallel with a game project and used them to evaluate the game. In addition, we have conducted the first validation of heuristics and evaluated five mobile games. The results indicated that these heuristics are useful in identifying playability problems in mobile games. Game usability and mobility problems are quite easy to identify since the procedure is very similar to evaluation of utility software. Playability problems related to gameplay are more difficult to identify, but gameplay heuristics helped evaluators to focus on important aspects in the gameplay.

7. ACKNOWLEDGEMENTS

We thank our colleagues Jussi Holopainen, Greg Costikyan, Jouka Mattila, Ciaran Harris, and Juha Arrasvuori for reviewing and commenting on the playability heuristics and Laura Ermi and Anu Jäppinen from University of Tampere for particularly useful comments to the model.

8. REFERENCES

[1] Björk S., Holopainen J., *Patterns in Game Design*, Charles River Media, Hingham, MA, 2005.

[2] Csikszentmihalyi M., *Flow: The psychology of Optimal Experience*, Harper Perennial, MA, 2005.

[3] Desurvire H., Caplan M., Toth J.A., *Using Heuristics to Evaluate the Playability of Games*. In Proceedings of Computer-Human Interaction 2004, pp. 1509-1512.

[4] Falstein N., *Better by Design*. Column Series in Game Developers Magazine, 2004

[5] Federoff M., *Heuristics and Usability Guidelines for the Creation and Evaluation of FUN in Video Games*. Thesis at the University Graduate School of Indiana University, Dec 2002.

[6] Idean Research, *Series 60 Developer Platform 2.0: Usability Guidelines for Symbian C++ Games*, Forum Nokia, Available at www.forum.nokia.com

[7] International Organization for Standardization (ISO) 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals, Part 11: Guidance on Usability, 1998

[8] Koivisto E.M.I, Korhonen H., *Mobile Game Playability Heuristics*, 2006, Available at www.forum.nokia.com

[9] Järvinen, A., Heliö, S., and Mäyrä, F., *Communication and Community in Digital Entertainment Services – Prestudy Research Report.* University of Tampere, 2002. Available at tampub.uta.fi/tup/951-44-5432-4.pdf.

[10] Malone, T.W., *Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games.* In Proceedings of the 3rd ACM SIGSMALL Symposium and the first SIGPC Symposium 1980, pp. 162-169.

[11] Muller, M.J., McClard, A., Bell, B., Dooley, S., Meiskey, L., Meskill, J.A., Sparks, R., Tellam, D. *Validating and Extension to Participatory Heuristic Evaluation: Quality of Work and Quality of Work Life.* In Proceedings of Computer Human Interaction 1995, pp. 115-116.

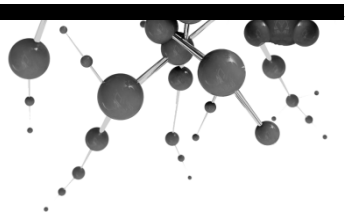
[12] Mulligan, J., and Patrovsky, B. *Developing Online Games.* New Riders, Indiana, 2003. pp 137.

[13] Nielsen, J., *Enhancing the Explanatory Power of Usability Heuristics*. In proceedings of Computer Human Interaction 1994, pp 152-158.

[14] Nielsen, J., *Usability Engineering*, Academic Press, London, 1993.

[15] Nielsen, J., Molich R. *Heuristic Evaluation of User Interfaces*. In Proceedings of Computer Human Interaction 1990, pp. 249-256.

[16] Rollings A., Adams E., *On Game Design*, New Riders, Indiana, 2003.



Paper II

Korhonen, H. & Koivisto, E. M. I. (2007). Playability Heuristics for Mobile Multi-player Games. In *Proceedings of the International Conference on Digital Interactive Media in Entertainment and Arts (DIMEA 2007, Perth, Australia)*, 28–35. New York, NY, USA: ACM. http://doi.acm.org/10.1145/1306813.1306828

Playability Heuristics for Mobile Multi-player Games

Hannu Korhonen
Nokia Research Center
Visiokatu 1, 33720 Tampere, Finland
hannu.j.korhonen@nokia.com

Elina M.I. Koivisto
Nokia Research Center
Itämerenkatu, 01800 Helsinki, Finland
elina.m.koivisto@nokia.com

1. ABSTRACT
Multi-player games are engaging because of social interaction and competing with real players. Currently, many digital games are multi-player or have multi-player features. When evaluating the playability of multi-player games, we need to consider player-to-player interaction. In this paper, we introduce and describe playability heuristics for mobile multi-player games that complement the playability heuristic model that we have presented earlier. Multi-player heuristics have been developed by identifying playability problems in game evaluations and comparing these findings to game design guidelines and other studies found in the literature. The results indicate that the presented multi-player heuristics correspond to issues that are addressed to be important in multi-player games.

Categories and Subject Descriptors
J.m. Computer applications. Miscellaneous

General terms
Design

Keywords
Playability, Multi-player Games, Mobile, Heuristic Evaluation

2. INTRODUCTION
Multi-player games are often considered to be more interesting and challenging than single-player games. The reason for this is player-to-player interaction; playing a game against another player instead of artificial intelligence (AI) is typically more unpredictable and enjoyable. In addition, the game sessions can be socially pleasant.

In a multi-player game, two or more players play the game in a same game session. The players may play concurrently or the play sessions may be asynchronous. In the asynchronous play sessions, the players access the same game world, but not necessarily at the same time. Multi-player games can be divided into two categories: 1) online games and 2) proximity games. When considering how the game state is maintained from a game session to another, the multi-player games can be also divided into two categories: 1) persistent games and 2) non-persistent games. In the persistent games, the game state is typically maintained on game servers and the players connect to them with a game client.

In online games, players are connected to each other via the Internet or other kind of network technology, such as a peer-to-peer network. The players usually do not share the same physical space and they use their own device (a computer, a game console or a handheld device) for playing the game. In these games, the player population can range from a few to hundreds of thousands. For example, a popular Massively Multi-player Online Game (MMOG) World of Warcraft has 8 million registered players[1] and there have been 200,000 players playing the game simultaneously[2].

A single device can also be used for playing a multi-player game. There are three common ways of doing this. First, players can use game controllers to connect to a centralized computer or a game console to remotely control the game. Second, the players can use one game controller and take turns to play the game. This is typically called "hot seating". Third, the players can use one gaming device and share the keypad or the keyboard while playing the game simultaneously.

The objective of this study is to find out issues that affect the playability of mobile multi-player games and convert them to playability heuristics. These heuristics are similar to usability heuristics that are commonly used for evaluating productive software with the heuristic evaluation method.

Although there are various kinds of multi-player games, in this study, we explore online games that are played with mobile phones. The mobile phone has its specific features that need to be taken into account when designing games for it – some offering new possibilities and some challenges (see e.g. [12]). Even if these heuristics have been designed for games that are played with mobile phones, most of the issues that have been mentioned hold true in non-mobile games as well.

The paper is organized in a following way. First, we present our playability heuristic model. We review the related work and propose an initial multi-player mobile game playability heuristics based on that. Then we summarize e results of playability evaluations of three commercial mobile games. Based on these results, we add two new heuristics in our initial set of heuristics. After that, we have conducted an informal review of four commercial non-mobile multi-player games, to see how well our refined set of heuristics can be used for evaluating multi-player games in general.

[1] http://www.blizzard.com/press/070111.shtml

[2] Although, it must be noted that all of these players did not share a common game world, since World of Warcraft consists of several instances of the game world.

DIMEA '07 Perth, Western Australia

3. Playability Heuristic Model

We have previously introduced a playability heuristic model that can be used for evaluating mobile games with the expert evaluation method [12, 14]. Currently, the model consists of three modules: *Gameplay*, *Game Usability*, and *Mobility* (Figure 1). These modules form the core of the model and they are common for all mobile games.

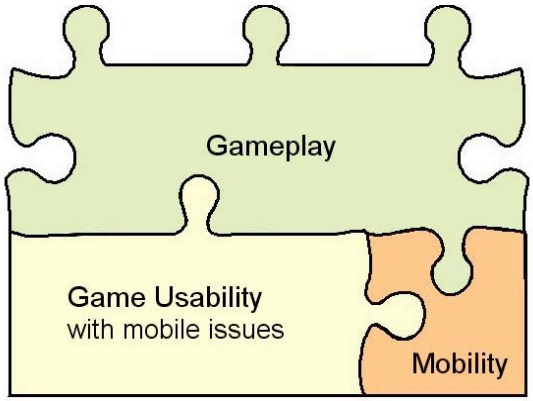


Figure 1. Playability Heuristic Core Model

Game Usability covers game controls and interface through which the player interacts with the game. As a general rule, the game interface should allow the player to control the game fluently and present all necessary information about the game status and possible actions [12]. *Mobility* concerns, as the name suggests, the mobile aspects of the game. The mobile phone has specific features that need to be taken into account when designing games for it. Mobile games can be played anywhere and anytime. The game design should assimilate this freedom into the game experience as well [12]. These two modules are the foundation of the model and they have many similarities to traditional usability heuristics. However, they have been tailored here for game evaluations.

Evaluating games with only traditional usability heuristics would leave many important aspects unprocessed [6]. Once playability problems covered by the *Game Usability* and *Mobility* modules are identified and fixed, evaluators can focus on gameplay, which is "the heart of the game". *Gameplay* heuristics deal with issues that arise when a player interacts with the game mechanics [12].

The modularity makes the heuristic model flexible [14]. The evaluators can focus solely on certain aspects of the design and use one or two modules at time. For example, the *Mobility* module is needed when evaluating mobile games and it can be ignored with other game types.

In our previous game evaluations, we noticed that there are some playability problems that current model is not capable of describing them [14]. These problems were closely related to gameplay, but they appear only in multi-player games. This result indicated that we need supplemental heuristics when evaluating multi-player games and they should be presented in a separate module.

4. RELATED WORK

We reviewed multi-player game studies to find initial set of issues that affect playability of multi-player games. In the following, we discus these issues and in the end propose an initial set of multi-player gameplay heuristics.

Finding other players is a fundamental issue in any online multi-player game [11]. Before a player can communicate and collaborate with other players, he or she needs to find them. In the online game worlds, the player does not necessarily know other players beforehand and making new contacts and friends is important [20].

The importance of player-to-player communication is stressed in the game design literature [7, 15, 19, 20]. If the players are not able to communicate, it is unlikely that the game attracts a community [11]. Communities are considered to be an essential part of multi-player gaming, particularly in Massively Multi-player Online Games (MMOGs) [15]. Cornett states that lack of communication between new and other players was probably the most significant obstacle for new players to get into the game that they studied [5].

Collaboration is related with communication, since it often gives the players a reason to communicate. Collaboration is also highlighted in many multi-player games by encouraging or even enforcing the players to team up with other players by providing quests and tasks that require more than one player. Multi-player games often provide character classes or types that support or complement each other. Zagal et al. list social interaction, competition and cooperation as main design guidelines for a multi-player game [21]. Even if collaboration is commonly encouraged, in the case of MMOGs, it needs to be noted, that some of the players may want to play multi-player games sometimes alone. According to Blizzard [18] one of the reasons for the success of Blizzard's World of Warcraft game is that the player can solo, i.e. play alone, their characters to the maximum level.

Some researchers have raised a concern about antisocial behavior in multi-user computer mediated systems [3]. The problem exists in multi-player games as well. The players who become targets of anti-social behavior may quit playing the game [10] or cause increased customer service expenses. Deviant behavior in multi-player games can be defined as cheating [1], exploiting, hacking, and grief play [8, 15].

The network latency and its effects on gaming experience is a well-studied subject. There are numerous technical solutions how to minimize network latency (see for example [4, 9]). Pantel and Wolf have studied effect of network delay in real-time car racing multi-player games [17]. They concluded that the delay should not be more than 100ms, otherwise it affects player's gaming experience and makes controlling the game character more difficult. The increased delay will also create frustration and players may leave the game because of this [4][16].

Common trends in research indicate that there are six issues that affect playability of multi-player games and they should be included in playability heuristics. The proposed playability heuristics are:

H1: The game supports communication [5,7,11,15, 19,20]

H2: There are reasons to communicate [20,21]

H3: The game helps the player to find other players and game instances [5,11,20]

H4: The game supports groups and communities [15, 21]

H5: The design minimizes deviant behavior [1,3,10]

H6: The design hides the effects of network [4,9,16,17]

5. PLAYABILITY PROBLEMS IN MULTI-PLAYER GAMES

We evaluated three commercial mobile games to identify playability problems in multi-player mobile games and confirm our findings in the literature study. The games represented different game genres, since we wanted to see if similar playability problems emerge regardless of the style. Target player groups and devices are listed in Table 1. We evaluated the first (alpha) versions of the games, which means that most of the features were implemented. The evaluation methods that were used are discussed later.

	Game Style	Target Player	Device
A	Combat	20+, Male	A portable gaming device
B	Adventure	18+, Male	Smart phone
C	Puzzle	15+, Neutral	Smart phone

Table 1. Characteristics of evaluated multi-player games

Game A was a combat game where the player could duel other players in a multi-player game world. The game could be played in a single-player mode and multi-player mode. Although the single-player mode was interesting and educating, the real challenge was in the duels against other players. The single-player mode was used for practicing various tactics and developing teams for duels. We evaluated the game by using the heuristic evaluation method and playing it in both proximity and remote locations. This method highlighted playability problems that the game had in the multi-player mode.

Game B was an adventure game where the player explored the game world and solved missions. It was a massively multi-player game, although in the testing we had a limited number of players. The players could play single player missions or duel and trade with others. Some parts of the game, like the duel, were synchronous, and some, like the auction system, asynchronous. The game was evaluated several times. First, we conducted a heuristic evaluation of the game in order to remove the most obvious playability problems. Then we evaluated the game using a field-testing method. About 30 players played the game over three weeks and players' comments were collected and analyzed.

Game C was a puzzle game and the player's objective was to keep the game area clear. The game consisted of single-player and multi-player modes. The multi-player mode was similar as the single-player mode, but it was played against other players. We applied two evaluation methods for evaluating the playability of the game. First, we evaluated the game with the heuristic evaluation method. The heuristic evaluations consisted of two iterative sessions, a very early one, and another conducted later when the game was more complete. After the heuristic evaluation, we arranged a field testing with about 20 players participating in the test over one week.

It must be noted, that the test setup has affected the multi-player features since the number of the players in the testing situations was significantly smaller than what it typically would be in a successful already-released multi-player game. However, the test setup highlighted even better the issues that can be problematic when there are not enough players around when a player would like to play with someone.

In the following, we list and analyze the playability problems found in these evaluations. The potential problems in the games were reported to the game producer and developer and most of them were fixed[3].

5.1 Visibility of Other Players

If the players cannot see other players who are present in the multi-player game, playing the game can feel lonely. This was noticed in the each game evaluation that was done in this study. Even though there could be tens of players playing the game, the players did not always notice that there were others around. The players were interested in their friends' status and whether they are online at the moment. A buddy list was included in each game, but in game A, the list was sorted chronologically and recent friends were hard to find. In game B, the whole buddy list[4] was stored on a game server and updating the list was very slow due to the network latency. In game C, the player who was added to the buddy list needed to approve it first (which was probably done for preventing anti-social behavior in the game).

None of these games included additional information about the players or game characters, such as biographies or pictures. Having additional information about the other players could work as an "ice breaker" and help starting conversations. In game B, the game character contained statistical information about battles that the player had played. Unfortunately, results from both single-player and multi-player modes were combined. This made the high-score lists less informative because the scores were more related to the time that the players had used for playing the game, and not as much their skill.

Game A did not provide a common space where players could meet, chat and spend time together. Game B had areas where the players could have socialized, but unfortunately, they could be accessed only as single-player instances. Game C had game rooms where the player could meet other players, but they were not working well. The only activities that players could do there was chatting and waiting. All the players were listed there, but those who were battling seemed to be inactive to others because they could not send messages nor accept new challenges. The players ended up creating a lot of game rooms and the players spent a lot of time trying to find a room where someone would be active and willing to play a game together.

Basically the only method for finding new players in game A and C was to announce an open challenge and hope that somebody would accept it. Unfortunately, this method was considered tiresome because the probability that someone would actually notice the challenge was very small.

In game B, it was easy to find other players because all the players who were online were listed. The player could either send a message or challenge the player. However, a direct challenge was sometimes considered to be obtrusive because the other player might be in the middle of other mission in the single-player mode

[3]. In the game playability testing process that we have used, it is always the game producer's and development team's responsibility to choose which problems need to be fixed. Sometimes, the problems are related to important game features that should not be changed, and sometimes, they just cannot be fixed, due to technical or resource limitations.

[4] Instead of only online status

and the challenger could not attach a personal message to the challenge.

Field testing of game C introduced a new problem that was not found in the heuristic evaluations. Players from various time zones participated the testing. There were players from different time zones as they were from Europe, Northern America, and Asia. The time difference was several hours, which mean that players needed to be very flexible if they wanted to play together.

5.2 Communication with Other Players

The most critical playability problem in all games was communication. It was not possible to communicate while battling against other players in any of the games that were evaluated. This problem was not only related to the game design, but also to the fact that the mobile phones typically have small key pads that cannot be easily used simultaneously for controlling a game and sending messages. The lack of communication became evident when the games were played so that two players were in the same room. During those test sessions, the players commented performance, talked about overall situation of the duel or just chatted. There was a lot of communication and it made the game more enjoyable. Our test players often commented that it would have been nice to comment on the game during the game or afterwards.

Due to the lack of communication possibilities in the game, the players sometimes sent text messages to talk about the game. Text messaging was considered to be inconvenient communication method, especially when players tried to make contacts with new players and did not know their phone numbers.

There were many situations in which a player needed to say something to another player. For example, in game A before a duel could start, both players needed to select a team or create a new one for a duel. Sometimes this took a long time and it would have been polite to inform the opponent about the delay. However, the game did not provide any means for communication and the other player just saw a notification "Waiting for a player" without knowing the real reason for delay. In the case of game C, it was noted that it would be good to be able to send a message that would explain the reason for rejecting the challenge. The messages could be free text or predefined messages like "You are too experienced for me", "Sorry I'm busy at the moment, try me later". The predefined messages would have made the messaging faster.

5.3 Interacting with other players

Even though the games that were evaluated were multi-player games, the players did not interact much with other players. There was no collaboration between the other players. One of the reasons for this was that the games were rather small and adding collaborative features would have expanded the size of the game project significantly. In each game, player-to-player interaction was mostly limited to duels. Some trading aspects were included in the game A and B.

It was sometimes difficult to notice when the opponent was another player or artificial intelligence. Strategies and tactics were basically the same in both cases. This was probably caused by limited set of actions that were available in the duel. Also, the lack of possibilities of player-to-player communication also contributed to this issue. In game C, the gameplay was so intensive and hectic that players could not really follow what the other player was doing.

5.4 Comparing Oneself to Others

Finding appropriate opponents for duels was considered to be difficult in each game that was evaluated. The players would have wanted either high score lists or some other means to compare their performance against others. Proceeding in the game was visible in the player's avatar in game B. The more advanced and visually bigger avatar the player had, the harder it to duel with. Game A did not show enough information about players' skills or experience when receiving a duel challenge. In game C, it was possible to get a hint about other players by looking at high score lists in the single-player mode, since best scores were uploaded when connecting to the multi-player mode.

5.5 Supporting Groups and Communities

Player-created groups or communities were not supported in any of the games. Groups are usually related to collaboration in the game world and there was not much need for collaboration in the games that were tested. However, games A and B would have had potential for player communities. Game B assigned a faction for the player when he joined the game, but the players could not choose or change it. Also, the meaning of belonging to a faction was not clear for the players (it did not affect the gameplay). In game C, players played individually, since the game was one-to-one puzzle competition.

5.6 Network Latency

Players experienced high levels of network latency in game B and it made playing the game difficult. Since all game data was stored on a server, the mobile device needed to request information frequently. During the evaluation, requesting status of other players or vendor items in game shops took a very long time. Moreover, there were disconnections, which disrupted duels. Even though this part of the game was not yet fully finished, it demonstrated what will happen if there are problems with the network connection. Game C suffered from connection problems to game servers and sometimes the players just disappeared from game rooms.

5.7 Deviant Behavior

Some players in game C reported that cheating occurred in the game. The game was designed so that if the network connection was lost, the challenge was a draw. Some players were reported to use this feature on purpose and disconnected from the game if they were losing the challenge. The game counted wins and losses in the rankings, but the draws were discarded.

6. MULTI-PLAYER MOBILE GAME HEURISTICS

Our game evaluations identified playability problems related to multi-player games that should be addressed before launching the games. Potential problems related to all of the six multi-player heuristics, which we proposed based on our literature review, were found in the evaluations. Also, we identified two new heuristics: "*The game provides information about other players*" and "*The design overcomes lack of players and enables soloing*".

Based on our multi-player game evaluations and literature review, we have formed the following heuristics for mobile multi-player games Table 2. The heuristics are described more in detail in the following sections.

As any heuristics, these are only guidelines and there might be a valid design reason to break a heuristic in a game. A good

example of this, that even if it is a good guideline to support communication, it could be sometimes useful to limit communication to create interesting game mechanics [2]. Also, the importance of the heuristics depends on the scope and focus of the game. For instance, it might not be needed to include complex community features in a small simple puzzle game.

MP1	The game supports communication
MP2	There are reasons to communicate
MP3	The game supports groups and communities
MP4	The game helps the player to find other players and game instances
MP5	The game provides information about other players
MP6	The design overcomes lack of players and enables soloing
MP7	The design minimizes deviant behavior
MP8	The design hides the effects of the network

Table 2 Multi-player Heuristics

MP1. The Game Supports Communication

The players should be aware of other players and be able to interact with them in a way or another in every multi-player game. Otherwise the game does not have a reason for being a multi-player game. When the players communicate – either directly by talking or indirectly by acting - they are more likely to form communities and make friends. In the more complex multi-player games, the friends in the game are usually a major reason for the players to keep playing the game [10].

Depending on the game, communication channels for different purposes need to be supported. Communication can be both synchronous and asynchronous. Asynchronous communication is often useful in games that last longer than one play session. Synchronous communication is useful during the play sessions when the communication is related to game events. The players should also be able to choose with whom they like to communicate and ignore other players if necessary.

Sometimes it is feasible to limit the communication, in order to reduce spam or create interesting game mechanics [2]. Hiding tactical information can be used for creating a need for collaboration between the players. Limitations can be needed for keeping the amount of messages that a player receives reasonable. The communication can be possible only between the players who are nearby in the game world.

Mobile phones provide two means of communications: voice and text. Even if mobile phones are designed for communicating with other people, typing text with a small keypad and viewing information on a small screen during a play session can be cumbersome. Voice communication would be an ideal solution and it is faster than typing. However, implementing a voice chat during the gameplay can be technically difficult, although, possible. A mobile game Pathway to Glory[5] includes sending short voice-messages to other players while playing the game.

[5] www.pathwaytoglory.com

MP2. There Are Reasons to Communicate

A game should provide meaningful issues for the players to discuss about. When the player is able to see how other players are doing in the game, it will generate discussion topics. In a very simple game it may be interesting to discuss about the game tactics. More complex games can provide events, difficult boss monsters that require collaboration, interesting game objects, or puzzles to give the players more reasons to communicate. Common interests make it often more enjoyable to discuss with other people.

MP3. The Game Supports Groups and Communities

Particularly the more complex multi-player games benefit of supporting groups and communities. In a small, simple, mobile game it might not be feasible to implement the community-related features in the game, but it still could be offer something in other platforms instead (e.g. in a web page) or include features that could encourage the players to form communities and talk about the game outside the game. The players who feel that they belong to a community are more likely to keep playing the game [19].

Players should be able to collaborate with each other and form groups when necessary. It can be useful to provide the players means for organizing temporary groups or longer-lasting communities. The features that support this are, for instance, associating players with certain roles or providing an enclosed communication channel for members. In groups, it can be useful to support observing group members and seeing their current status [11].

MP4. The Game Helps the Player to Find Other Players and Game Instances

Multi-player games are usually played with other players. Players should sense that there are others around. If players do not know the others advance, the game should provide means for meeting and getting to know new players in the game world. The game design can provide a search feature.

If the game design includes game instances, the player should be able to easily find and join them. It can be useful to allow the non-players to be spectators before joining a group or the game.

MP5. The Game Provides Information about Other Players

Knowing some information about the other players is important in multi-player games. This does not only increase bond with other players in the game, but it also makes the players aware of the presence of other players. In competitive situations, finding a proper opponent is an important aspect of the game. In collaborative situations, the player needs to know other players' location and other properties, e.g., whether they need assistance. Information about the players' past performance can be also used to reduce deviant behavior.

MP6. The Design Overcomes Lack of Players and Enables Soloing

Many online games have a "critical-mass problem", which means that there are not always enough players in the game world to make the gameplay meaningful or waiting other players to show up can take some time. The game design should take this into account and provide the players something meaningful to do

while there are not enough other players around. For example, the player could compete against AI while waiting for other players.

Furthermore, the more complex multi-player games should provide content for solo players. Depending on the playing style, current context or mood, the player may prefer play alone. The player should be able to do this even if it is a multi-player game world [18].

It can be useful to allow players to distribute a client version of a game to the other players who do not own it. This improves greatly the possibility of finding other players to play with. If the game can be played only with those who own a version of the game, this feature can be also used as a tool for demonstrating and marketing the game.

MP7. The Design Minimizes Deviant Behavior

When players are in close proximity or know each other well, the probability for bad behavior is smaller. Online multi-player games often facilitate deviant behavior. Examples of such behavior in are cheating, exploiting, hacking, and grief play.

Cheating can be defined as "An action by a player that violates the rules of the game 'as written' or commonly understood" [4]. An exploit can be understood as a technique for cheating in the game. Hacking, in this context, means an act of creating an exploit. Grief play can be defined as "Play styles that disrupt another player's gaming experience, usually with specific intention to" [6].

Often, preventing grief play requires restricting player-to-player interaction. The balance between player-to-player interaction and preventing grief play should be carefully considered. Setting too strict restrictions may end up in dull gameplay.

MP8. The Game Design Hides the Effects of the Network

There are three major network-related issues that need to be taken into account in online games: latency, disconnections, and pricing of data traffic. The latency can disrupt the gameplay and cause delays to interaction. Mobile games can be designed so that they hide the latency [16]. If a player gets disconnected from the game, it needs to be handled gracefully [13]. The amount of data to be transferred between server and mobile devices can become a barrier for playing the game. Currently mobile network is still slower than broadband connections and players usually need to pay for every kilobyte of transferred data.

7. EXPERIMENTATION OF THE PROPOSED SET OF HEURISTICS

In order to try out our new set of mobile multi-player game playability heuristics, we conducted an informal and brief playability evaluation of commercial multi-player games. It must be noted that further studies are needed to formally validate the heuristics. However, our study shows what kinds of issues in the games that we selected were related to the heuristics. It also shows that the multi-player mobile game playability heuristics can be applied to non-mobile games as well.

We selected PC games from various game genres for the evaluation. The selected games were *Star Wars Galaxies* (MMORPG), *World of Warcraft* (MMORPG), *F1 Challenge '99-'02* (Racing Simulator), *TrackMania Nations* (Arcade Racing), *Call of Duty 2* (First Person Shooter), and *Civilization IV* (Command and Control). Next, we go through the observations related to the multi-player features in these games. The games are referenced as following abbreviations SWG, WoW, F1C, TMN, CoD, and CIV respectively.

7.1 Communication in Games

All the games that were evaluated include a chat feature. Players are able to send public and private messages. In PC games, the keyboard enables convenient writing. Usually there is a separate area in the user interface that shows players' discussion. In SWG and WoW it is possible to have several chat screens open at the same time, which makes it possible to follow multiple discussions.

One playability problem related to communication was found in CIV and F1C. In these games, the chat messages are visible only for a short period of time and there was no chat log available. It is possible, that some messages are not noticed during the hectic gameplay.

7.2 Reasons to Communicate

None of the games is lacking of finding reasons for communication. Due to convenient typing with the keyboard, we noted that the players communicate frequently about the game and often also about non-game-related issues.

7.3 Support for Groups and Communities

All of the evaluated games, except one, support groups and communities[6] at least in some extent. In SWG and WoW, the players can create temporary groups for missions and the games support forming communities. In CoD, the players can choose their favorite army and in F1C their team.

The design of TMN does not explicitly support or encourage the players to form communities. However, the players form communities themselves. They often indicate favorite game server or community in their nickname. Usually it is done in format of "[community] nickname". In SWG and WoW community features are handled automatically

7.4 Finding Other Players and Game Instances

In SWG and WoW, a player can find other players just by wandering in the game world. Both games also include support for finding other players, e.g. in form of a looking-for-group flag. If the player wants to play with real-life friends, one problem is that there are several game servers and the player needs to know on which server to join. In these games, if the player creates a character in one game server, it cannot be easily moved to another one later. In the other games that we studied, the player can freely move between servers to find players to play with. All the games have buddy lists, which helps to seeing the current status of friends.

7.5 Providing Information about Other Players

SWG and WoW are good examples of games that provide information about other players and the players can customize their game characters. In CIV, the information about the other players is hidden on purpose and the players will find out more details while they progress in the game.

[6] In MMOGs, a common example of a player community is a guild.

The lack of information about the other players could be considered as a playability problem in TMN. The game does not provide much information on how other players are doing until the race is finished. This is totally the opposite compared to F1C, which is similar kind of game. In F1C, the players know all the time what is the gap between the players who are behind and ahead. This will create tension to the race and push the player to try even harder.

CoD provides information about other players just enough to avoid friendly fire. This is common in the other games in this genre (First Person Shooters, FPS) and it seems to be usually enough for the players.

7.6 Overcoming Lack of Players

SWG, WoW, TMN provide game content also for solo players. In SWG and WoW, the player can choose various missions and explore the game world alone. TMN allows the player to enter the game instances alone and practice tracks if there are no other players available. Other players can join the race at any time. CoD, F1C and CIV have separate single-player modes.

7.7 Minimizing Deviant Behavior

Grief play and other forms of deviant behavior are commonly acknowledged problems in MMOGs and SWG and WoW are no exceptions to this. The games try to overcome this problem with various solutions, e.g. in World of Warcraft, the groups can set up specific rules how the loot of a killed monster is shared between the group members.

The players cheating by disconnecting from the game before they lose was identified as a problem in Game 3 that we evaluated earlier. TMN implements a solution to this problem by counting in also the results of the players who have disconnected before the end of the race. However, the solution also makes disconnections that do not happen on purpose more annoying for the players. This demonstrates how preventing deviant behavior and providing enjoyable gameplay may be contradicting and the designers sometimes need to compromise.

Tweaking units and game characters can be sometimes considered as a problem and sometimes as part of the gameplay. In CIV, the players can modify game units in order to make the game more suitable for them. In the multiplayer mode, it can be a problem if players have different versions of units.

7.8 Hiding the Effects of the Network

All the games that we studied have a pretty good client-server implementation. The worst problems occur in fast-paced games. In F1C and TMN, if network latency occurs, the players lose the race and they have to wait for the next race. On the other hand, TMN provides very short races (approximately five minutes) and the players are able to attempt to improve their time on the track.

8. DISCUSSION

Our studies revealed that that the proposed set of multi-player mobile-game heuristics addresses quite well the issues that needs to be taken into account when evaluating multi-player games. The study on non-mobile multi-player games also showed that the heuristics are relevant in the case of other multi-player games as well. The mobile game playability studies showed that if the multi-player aspects are not implemented properly, the game is less enjoyable. Our mobile game evaluations also indicated that when multi-player games are evaluated, the requirement of having a critical mass of players must be taken into account. This means that in order to experience multi-player features in the game, there must be enough players in playing the game. When considering the test setup, we noted that this is particularly important in field testing where players are free to play the game whenever they want.

When conducting evaluations of multi-player games playability experts should consider which heuristics are relevant for the game that is being evaluated. Because there are several types of multi-player games, some heuristics may not be applicable. For example the eighth multi-player heuristic about network is feasible only for online games.

In the future studies, we will explore if the multi-player heuristics cover all critical aspects. The literature review and game evaluations suggest that we were able to catch at least the most obvious issues.

9. CONCLUSIONS

In this paper, we explored playability problems in mobile multi-player games and expanded our playability heuristic model with a new module. The multi-player heuristics were formulated based on findings from the literature review and supplemented with findings from playability evaluations that we conducted with mobile multi-player games. The result of the study was that we identified eight playability heuristics related to multi-player games. These heuristics are related to communication, collaboration, deviant behavior, amount of players, visibility of other players, and social interaction in groups and communities. In addition, there is a heuristic related to network connection, which is an important part in any online game, particularly when considering mobile games.

Multi-player heuristics complement the playability heuristic model and increase usefulness of the model since it can be applied in evaluations of a broader range of mobile games.

In the future, we need to further validate these multi-player heuristics and evaluate a wider range of the multi-player games. There are also other game styles in mobile domain, which require own set of heuristics to be included to the model. Currently our model covers mobile games that resemble games from stationary game platforms. However, there are also unique games styles for mobile devices that probably need their own set of heuristics in evaluations. Examples of these games styles are location-based games and context-aware games.

10. ACKNOWLEDGEMENTS

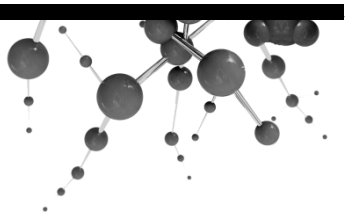
We would like to thank all the participants of our test sessions and the game producers and developers of the game titles. Without these people the evaluations would not have been possible. Our special thanks go to Mark Ollila who has provided useful comments on this paper.

11. REFERENCES

1. Baughman N.E., Levine B.N., Cheat-proof Layout for Centralized and Distributed Online Games. In *Proceedings of* the *Twentieth IEEE Computer and Communication Society INFOCOM Conference*, IEEE, 2001, pp. 104-113.
2. Björk, S., Holopainen, J. *Patterns in Game Design*. Charles River Media, Hingham, MA, 2005.
3. Bruckman A., Curtis P., Figallo C., Laurel B., Approaches to Managing Deviant Behaviour in Virtual Communities. In

Proceedings of ACM CHI 1994, ACM Press (1994), pp.183-184.

4. Chen K-T., Huang C-Y., Huang P., Lei C-L., An Empirical Evaluation of TCP Performance in Online Games. In *Proceedings* of *ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ACM Press (2006).
5. Cornett S., The Usability of Massively Multi-player Online Roleplaying Games: Designing for New Users. In *Proceedings of ACM CHI 2004*, ACM Press (2004) pp. 703-710.
6. Federoff M. Heuristics and Usability Guidelines for the Creation and Evaluation of FUN in Video Games. *Thesis at the University Graduate School of Indiana University*, 2002.
7. Friedl M., Online Game Interactivity Theory, *Charles River Media*, 2003.
8. Foo, C.Y., Koivisto, E.M.I. Defining Grief Play in MMORPGs: Player and Developer Perceptions. In *Proceedings of ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ACM Press (2004), pp.245-250.
9. Gautier L., Diot C., Design and Evaluation of MiMaze, a Multi-player Game on the Internet. In *Proceedings of IEEE Multimedia Systems Conference*, 1998, pp 233-236.
10. Jensen C., Davis J. Farnham S., Finding Others Online: Reputation Systems for Social Online Spaces. In *Proceedings of ACM CHI 2002*, ACM Press (2002), pp. 447-454.
11. Koivisto, E.M.I. Supporting Communities in MMORPGs by Game Design. In *Proceedings of Level Up! Games Research Conference*, 2003.
12. Koivisto, E.M.I. and Palm, T. Iterative Design of Mobile Games. In *Proceedings of Game Design and Technology Workshop*, Liverpool, UK, 2005.
13. Koivisto, E.M.I and Korhonen, H. Mobile Game Playability Heuristics. *An article at Forum Nokia*. 2006. Available at http://www.forum.nokia.com.
14. Korhonen H., Koivisto E.M.I Playability Heuristic for Mobile Games. In *Proceedings of the 8th conference on Human-Computer Interaction with Mobile Devices and Services MobileHCI '06*, ACM Press (2006), pp. 9-16
15. Mulligan J., and Patrovsky B., *Developing Online Games*, New Riders Indiana, 2003
16. Palm, T., and Koivisto, E.M.I. Developing Online Mobile Games. *A Speech in Game Developers Conference Europe*, London, UK, 2004.
17. Pantel L., and Wolf L.C. On the Impact of Delay on Real-Time Multi-player Games, In *Proceedings of the 12th international workshop on Network and operating systems support for digital audio and video*, 2002, pp. 23-29.
18. Pardo, Rob. A key note speech on the design philosophy of World of Warcraft. *Austin Game Conference*, Austin, USA, 2006
19. Seay A.F, Jerome W.J., Lee K.S., and Kraut R.E., Project Massive: A Study of Online Gaming Communities, *Ext. Abstracts CHI 2004*, ACM Press (2004), pp 1421-1424.
20. Taylor T.L., *Play between Worlds: Exploring Online Game Culture*, MIT Press, 2005.
21. Zagal J.P., Nussbaum M., Rosas R., A model to Support the Design of Multi-player Games, *Presence: Teleoperators and Virtual Environments*, Vol. 9, No 5, MIT Press, 2000, pp. 448-462.



Paper III

Korhonen, H., Saarenpää, H. & Paavilainen, J. (2008). Pervasive Mobile Games – A New Mindset for Players and Developers. In Markopoulos, P., de Ruyter, B., Ijsselsteijn, W., & Rowland, D. (Eds.), *Proceedings of the International Conference on Fun and Games (Fun'n'Games 2008, Eindhoven, The Netherlands),* LNCS 5294, 21-32. Berlin Heidelberg, Germany: Springer-Verlag. doi: 10.1007/978-3-540-88322-7_3

Pervasive Mobile Games – A New Mindset for Players and Developers

Hannu Korhonen[1], Hannamari Saarenpää[2], and Janne Paavilainen[2]

[1] Nokia Research, P.O. Box 1000, 00045 Nokia Group, Finland
hannu.j.korhonen@nokia.com
[2] University of Tampere, Kanslerinrinne 1, 33014 Tampereen Yliopisto, Finland
{hannamari.saarenpaa,janne.paavilainen}@uta.fi

Abstract. Pervasive games are an emerging new game genre, which includes context information as an integral part of the game. These games differ from traditional games in that they expand spatio-temporal and social aspects of gaming. Mobile devices support this by enabling players to choose when and where a game is played. Designing pervasive games can be a challenging task, since it is not only limited to the virtual game world, but designers must consider information flow from the real world into the game world and vice versa. In this paper, we describe a user study with an experimental pervasive multiplayer mobile game. The objective was to understand how the players perceive pervasiveness in the game and what the crucial factors are in the design. Based on the results, we propose initial design guidelines and compare them to other design guidelines for the pervasive games.

Keywords: Mobile Game, Pervasive Game, Game Design, Design Guidelines, Context, Asynchronous gameplay.

1 Introduction

Pervasive games introduce a new emerging game genre that pushes the boundaries of the traditional games and enables new kinds of gaming experiences for players. One of the most exciting aspects in these games is that the context information is utilized to modify a game world or it is converted to game elements. In addition, gaming can be blended into the daily life and normal social situations of the players.

Pervasive gaming is a wide domain, which can consist of the real world games augmented with computing functionality, or virtual computer entertainment is brought back into the real world [12, 13]. Magerkurth et al. introduce several pervasive game genres such as smart toys, affective gaming, augmented tabletop or real world games, and location-aware games [12]. Even though it is not a comprehensive list of pervasive game genres, it gives a good overview to the broadness of the domain.

Our research focuses on pervasive games that are played with mobile devices. The mobile device is a good platform for pervasive games, since it is pervasive by its nature. It is capable of acquiring information about the current context and it can send

P. Markopoulos et al. (Eds.): Fun and Games 2008, LNCS 5294, pp. 21–32, 2008.

information (e.g. location) to a game system, which then defines the appropriate player context.

Designing pervasive mobile games is a challenging task as many new issues need to be taken into account in a design. As context information is a crucial element in these games, the designers should emphasize this aspect in the design as well. Moreover, the pervasive games are often played in environments inhabited by people who are not playing the game. The game design must ensure that the game does not disturb too much players' social interaction outside the gameworld or disrupt non-players' ongoing activities. Further, since the players may be distracted from their surroundings by focusing on the game at hand, they may become a hazard for themselves or others. Designing a game to avoid these problems is a key factor in acceptance of the pervasive games.

In this paper we describe a user study with an experimental pervasive mobile game. Our objective was to find out how players perceive pervasiveness and what issues are important in a pervasive mobile game design. The contribution of the paper is a list of initial design guidelines for pervasive mobile games.

2 Related Work

Pervasive games have been studied for several years, but there are only few design guidelines available for helping designers in their task. Eriksson et al. present design guidelines for socially adaptable games that are played in different social context than traditional games [5]. The guidelines are as follows: *support interruptability*, *allow multiple communication channels*, *consider ambiguity*, *design for external events*, *allow modes of play based on social roles*, *minimize social weight*, and *analyze intended player groups from several perspectives*. They highlight essential aspects in design and are focusing on how the game can be adapted to social environment in which the game takes place, but they are missing the role of context information.

Crossmedia games are one type of pervasive games that use multiple gaming devices in addition to media channels like TV and radio [14]. Ghellal et al. [6] discuss game design aspects that designers should take into account when designing pervasive crossmedia games. Many of the design challenges in crossmedia games are concentrated on using several devices at the same time, but pacing the game was also seen important. The initial design guidelines say that the game should support both active and passive participation as well as different temporal involvements. The problem with these guidelines is that they focus too much on the devices instead of players so they are not sufficient enough to be used for designing other pervasive games besides crossmedia games. For example these guidelines don't give any assistance on how one should design context sensitiveness.

Lankoski et al. describes a location-aware mixed reality game, *The Songs of North*, which was used to study design solutions for pervasive games [9]. One guideline they found was a support for communication since players can be present in the game at different times. They also discovered the importance of a player control which is related to the temporal gameplay; the player is not punished for not being present in the game or playing seldomly. Lankoski et al. did not present proper guidelines but

rather observations on important design issues. That is probably one reason why they seem to have missed issues like communication between the player and the server.

3 The Game Concept

Mythical: The Mobile Awakening[1] is an asynchronous slow-update multiplayer game where players access a magical world through their mobile phone. The magical world is divided into four factions (Dawn, Sun, Dusk, and Moon). The players gain experience and learn spells by completing rituals either alone or together with other players. The spells are then used in encounters to battle against AI opponents or other players (Figure 1B). The game content is based on folklore mysteries and local history for creating an exciting atmosphere.

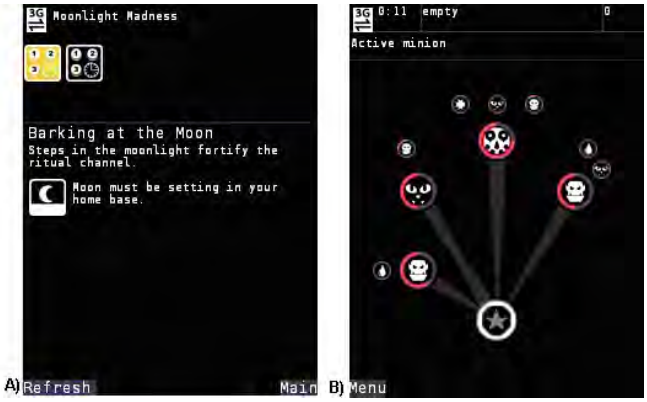


Fig. 1. Screenshots from the game. A) Context information sets conditions when the ritual component can be completed successfully. B) Slow update gameplay in the encounters.

The game features context-aware gameplay where the real world phenomena have an effect in the game world. Context information derived from the real world is used in the rituals where the reward of the ritual depends on how well the player has met the context conditions set initially (Figure 1A). There are three types of context information used: spatio-temporal, environmental and proximity. Spatio-temporal context is used in two ways: players select a home base from the predefined list and the game content and some environmental context information is then validated against information on that location. Time of the day is frequently used context information that defines when some rituals can be completed. Environmental context information is based on temperature, cloudiness and astronomy. Temperature is used in a breakpoint manner; some rituals require that the temperature is either above or below 0 degrees Celsius. Cloudiness has three possible options: clear, partly cloudy, and cloudy. Astronomy information is related to the Moon and Sun positions over the horizon and to the phases of the Moon. The proximity context is based on scanning Bluetooth devices. Rituals can require scanning either a specific or a given number of Bluetooth devices.

[1] http://www.mythicalmobile.com/

Asynchronous gameplay favors casual play style and Bogost has suggested that such feature could be the future of casual multiplayer gaming [3]. Slow update gameplay means that the game events are not continuous, but they happen in predefined intervals ranging from less than a minute to several hours.

The game was implemented upon MUPE[2] / Equip 2[3] platform and it is running on mobile phones supporting MIDP2.

4 User Study

To find out how pervasive features affect the gaming experience, we conducted a user study on the game on November 2007. Next we describe the procedure and the participants who took part in the study.

4.1 Participants

We had six participants, four females and two males, between 15 and 16 years. They were all regular video game players (approximately 10 to 25 hours per week). They played games mainly with a computer. Mobile games were not popular among the participants. Two participants played mobile games daily, but for others it was rare. None of the participants had any previous experience from pervasive games.

4.2 Procedure

The user study period lasted one week. The participants were instructed to play the game freely in their own time as much as they wanted to. Thereafter they were interviewed about their playing experiences. We conducted in-depth one-to-one interviews that lasted about 1.5 hours each. With these interviews we wanted to find out the overall opinion and attitudes towards the pervasive mobile game, and how the participants perceived context information that was embedded in the game design. The interviews revealed issues about how the implementation of these features had succeeded as well as opinions towards pervasive games in general. All interviews were audio recorded for later analysis. After the interview, the participants were also asked to fill in the background questionnaire, which revealed their gaming habits.

The interviews were processed in a way that all individual comments were written down to paper slips so each of them could be understood alone. After that we constructed an affinity diagram [2] from the approximately 700 comments we got from the interviews, combined related comments together, and gave a title for the group that described those comments (Figure 2). This method is well known in user-centered design projects for utility software, but in game evaluations it has not been used extensively as far as we know of.

We found this method to be time-consuming but useful; it forced us to think about each comment and what it actually meant. It also allowed us to see bigger entities as comments were not grouped based on the question, but according to their similarity of the topic. Results section describes some of the topics in more detail.

[2] http://www.mupe.org
[3] http://sourceforge.net/projects/equip/

Fig. 2. First level categorization of interview comments

5 Results

In this section we describe some of the interview results, which revealed that pervasive mobile games will undoubtedly change many existing practices, but on the other hand they also provide new possibilities for gaming.

5.1 Rethink Game and Play Sessions

One of the first new things for the participants was the concept of a game session. In traditional video games the game session has a clear beginning and an ending and play sessions are mostly continuous. In pervasive mobile games this is not anymore straightforward as the game sessions might be fragmented and the whole game session can last hours or even days. The participants commented that this new style requires some learning and they had some troubles at the beginning to get used to it.

U2: *"I didn't leave the encounters unfinished before exiting the game."*
U4: *"I felt mostly like I could leave the game at any time. Still I wanted to finish started rituals and battles."*
U6: *"I had to get used to the playing style that you can leave the game in the background and it continues."*

For other participants this playing style was intuitive and they had accustomed to fragmented play sessions and slow update game mechanism that enabled flexibility.

U5: *"I felt like I could leave the game at any time."*
U3: *"I just started the encounter and let it play alone. I did not have to do any preparations, just to make sure that the game can roll on its own."*

5.2 Leaving the Game Situation Active Makes the Player Curious about It

Having the game active on the game server after log out caused an interesting dilemma for the participants. Do they just let the game play alone or should they log in again to see that everything is going alright? The participants stated that they were interested in and wondering what was going on in the game while they were away.

U6: *"I came back to the game if I was even a slightly curious about it."*
U3: *"I thought the game in some extend even though I was not playing it."*
U4: *"I thought about the game at times when I wasn't playing it if I had just left the game aside and thought that I would continue it later."*

After returning to the game the participants checked the current game state and reviewed previous game events in encounters.

U5: *"I usually read the texts that appeared after the battle I had just finished."*
U4: *" I have used the log feature and I think it is useful."*
U3: *"I could follow game events in the encounters and observe what kinds of minions or spells the opponent used."*

However, reviewing the current game state was not satisfactory, since the participants could have missed something interesting or important and it was too late to influence to the game events anymore at that point. Some participants tried to overcome this shortcoming by accessing the game several times a day to be on the safe side.

U3: *"Sometimes when I entered duels they seem to be almost finished."*
U2: *"If there would have been more actions happening at times when I wasn't playing the game, then it would be more interesting and it would irritate if I would have missed something important."*
U6:*"I checked the game quite frequently and I did not leave it alone for an hour or more."*
U2: *"I checked the game at least two to five times per day."*

5.3 The Game Should Notify about Important Events

Based on the participants' comments the game should have kept the players aware of the game events. Three participants mentioned that the game could send SMS messages to a player when something important has happened in the game and the player is not online. This would reduce frequent logins to the game.

U5: *"I think the SMS feature would be fun. Then you don't have to access the game all the time."*
U4: *"The SMS feature would have make things easier when you wouldn't have to go into the game just to check out if something interesting has happened."*

The notification feature was built in to the game, but none of the participants had activated it. There were UI design problems related to this feature, but since it is a novel feature in mobile games, the participants did not expect to find it in the game.

5.4 Finding Proper Intervals for Slow Update Mode

The important question for slow update gameplay is how to define event intervals that support this gaming mode. The participants stated that they had two approaches for this. They used short intervals (30-45 seconds), in which events occurred almost in real time or considerably longer intervals, which left room for other activities.

U6: *"Sometimes it was good that the encounters lasted longer, because then I had a chance to do something else."*
U4: *"I think that the encounters are sometimes really slow but then again it makes it possible to do something else at the same time."*

The game event intervals between one to five minutes were considered problematic. They are too short for leaving the game unattended, but at the same time it was boring to wait for the next game event to take place.

U2: *"You cannot choose the wave length in every battle so it leaves quite vulnerable feeling when you have like 2 min waves and you cannot leave the game. You have to be in the game and see what happens."*

5.5 Gaming Blends into Other Activities

One of the key features in pervasive games is that gaming is not anymore an isolated activity, but the players can share their time with a game and other tasks at the same time. The participants confirmed that they were listening to music, reading emails, eating and doing other tasks while playing the game. Concurrence of gaming and other tasks is pretty much dependent on concentration level and time that is required to perform actions in the game. If that time is long, it will disturb other activities.

U5: *"This game did not change my daily routines. It was just an addition."*
U6: "It was really easy for me to concentrate on something else than playing the game even though it was running in the background".
U3: "I feel that it is not very distracting if I occasionally take a look what happens in the game world and it does not take very long time".

5.6 Utilizing Context Information Was Appealing

Context information made the game appealing because setting conditions and defining the success rate of actions was different than what the participants were familiar with the traditional video games.

U5: *"I liked those rituals the best where there were different times of the day."*
U3: *"Rituals seem reasonable, especially when you have a multi-component ritual, which can be completed only in certain times of the day".*

The challenge was to find the right context where the ritual could be completed. As the players cannot rely anymore on information provided only in the game world, they must also exploit other sources.

U3: *"I would probably try to figure out the current time of day by checking the outdoor or using a calendar."*

5.7 Sometimes It Is Not Possible to Play the Game

Finding out the right circumstances to play the game can be a challenging task. Therefore, it is tempting to play when there is a possibility. However, the players' social context or more urgent tasks will limit this possibility. The participants commented that they were thinking about their influence on others while playing the game because they might unintentionally be impolite towards them.

U5: *"There are those rituals tied to a certain time of the day. So as I belong to the School of Dusk, it is quite short time when the rituals are open and we usually have a class at that time. It is quite hard for me to play then."*
U4: *"I think the playing would be quite much depended on my daily schedule."*
U5: *"The others became impatient when I was playing the game and they had to wait for me to finish."*
U2: *"It is quite rude to play this game in certain situations where you really have to be involved."*

6 Design Implications

The user study results indicate that pervasiveness sets some new requirements to the game design. In this section we present implications that the designers should take into account when designing asynchronous pervasive mobile games. Sections 6.1 and 6.2 discuss about utilization of context information in a game. Sections 6.3 and 6.4 deal with issues related to asynchronous gameplay.

6.1 Perception of the Current Context

Pervasive mobile games typically use context information to define some conditions in the game world or to convert context information to a game element. One part of the fun in the pervasive games comes from discovering the correct context and playing the game when conditions are favorable for the player.

The difficulty in using context information in a game design is that the game system and a player may not have a mutual understanding of the current context. The game design should not rely on context information blindly because it may be inaccurate or unavailable due to various reasons [7]. Inaccuracy can be a result of measurement variation due to different locations. For example, temperature can fluctuate several degrees even within a close range.

A player may also be uncertain how the game system interprets current context information even though it would be otherwise clearly recognizable. For example, the game can utilize environment context information such as cloudiness in the game. Even though cloudy and partly cloudy skies are different by definition, from the player point of view they are alike compared to the clear sky.

Using rigid thresholds with context information should be avoided because information from different sources may vary slightly and thus becomes ambiguous. Therefore, the designer should decide whether similar context information is considered as same or treated as different in the design. The best option is to utilize only context information that is clearly discernible and unambiguous for the players.

Furthermore, the game should notify the player what the game system considers as current context. The difficult issue for the designers is how the context information should be revealed to the player without contaminating the gaming experience. In addition, there can be several contexts that need be informed simultaneously.

6.2 Equal Chance to Play

Utilization of context information in the game should be carefully designed and the designers need to make sure that all players have an equal possibility to access relevant information. The player progression should not be depended on context that may be unreachable. Furthermore, most context information is also dynamic and the designers cannot control changes (e.g. weather information), which may place players to unequal position.

In addition to uncontrollable changes in context, there can be player-related reasons why the player cannot access certain information. For example specific time of the day may not be suitable for the player due to other activities or the player's current social context does not allow gaming at that point. The player may not be able to travel to retrieve location specific information or some resources are not available at the player's current location. Our results confirm the previous study, which state that the lack of time and unfavorable location were reported to be the major obstacles for playing the game successfully [1]. Peitz et al. also report results of a pervasive game in which players were frustrated because there were not necessary game resources available at the player's location [15].

However, if some context is available only for some players, the game design should enable and emphasize collaboration between the players. With this solution all players may access context information equally. This will also increase player-to-player interaction in the game. Other possibility is that context information is time-independent and the player can use it when it becomes available. If certain context is not available, the player should always be able to do some other things that will require some other information. Finally, the game should provide meaningful things to do that do not require any context conditions to be fulfilled. Other alternative is that the player can complete the task, but the reward is determined based on how well the context conditions were met.

6.3 Adjustable Play Sessions

Players play pervasive games as time permits it. Gaming can be very sporadic and the game design should take this into account. Moreover, as the play sessions are blended into other activities it is possible that the player needs to leave the game without prior notice or play sessions are short by default. Peitz et al. argue that short play sessions will promote social adaptability of the pervasive mobile game [15].

The game should be designed so that it is possible to leave the game at any time without reducing the player's chance to win the game. Disconnections are likely to happen in pervasive games and the game design should manage these properly [4]. In the multiplayer games the design should consider what kinds of effects the abandonment of the player has on other players.

In asynchronous gameplay the player should be able to adjust the pace of the game and match it to available playing time. Sometimes it is preferable to play the game almost in real time while some other time a slower game pace is preferred as it gives more possibilities to lie back, and time to do other things at the same time. In this case the game requires minimal attention, but the player can still keep track on game events. Our study results indicate that the length of the short play session should be less than five minutes. These short play sessions allow the players to play the game and have some other activities at the same time without disturbing each other. Longer play sessions are preferred when the players have time to be invested for gaming.

6.4 Communication Outside Gameworld

Communication is a very important factor in multiplayer games. Players quite often create their own tools or utilize existing tools if the game design lacks required features (e.g. for communication) [8]. Even though the support for communication between the online players is a common feature in many games, asynchronous communication is gaining attention as well. Linner et al. introduce a framework, which enables rapid development of pervasive games that also supports offline communication [11]. The system buffers received messages and the player can see them on the next login. Lindley et al. also describes a pervasive mobile game that allows the players to receive alerts, although they are not actually playing the game (the game character is in a dormant mode) [10].

The participants in our study highlighted very clearly that they want to be aware of the game events when they are offline. This observation expands the current design practice because currently the communication stays inside the game world and the players will receive messages when they log in to the game next time.

From the interaction design point of view communication to the offline players is interesting because the game user interface expands outside the game and the game events are communicated by using some alternative communication channel. The designers should decide both the communication and presentation methods when providing information from the game world.

The designer can use direct or indirect communication methods. The benefit of the direct communication is that messages will be delivered immediately and the player can act without delays. The drawback of this solution is that it will generate costs to the game master because one has to use a messaging service such as SMS, MMS, or email. Indirect communication means that important game data is accessible through a medium, which does not require the player to log into the game world. One viable option is to use a web page, which displays the game events to the players. This solution does not create extra costs as the game server and the web server can be in the same local network and information is transferred internally. Indirect communication assumes that the players are active and will check the web page spontaneously. However, it is possible that the players will not get information in time.

Instead of using messaging services or web services for communication it is possible to build automatic update functionality to the game itself. In this solution the client connects to the game server at certain intervals and retrieves the recent game

events. Information is then displayed to the player. The player can view information without accessing the game, but if some actions are needed then the player has to log into the game. This solution creates frequent data traffic between the mobile device and the game server, and the player pays expenses of the traffic.

Presentation method depends on the communication method. SMS/MMS messages can have certain limits on how long messages can be, but for emails the length is not an issue. However, it should be noted that the emails are supposed to be read on a mobile device in order to achieve immediate information delivery. Textual information is in many cases sufficient, but images and graphics can make the messages more appealing.

Another important issue in a presentation method is to decide what game events are communicated to the player. In other words how many messages the player will receive when not logged in. The basic rule is that the player should be notified about anything that would cause actions when the player is actively playing the game. However, the number of the messages should not be overwhelming because it can easily lead to negligence and may disturb player's other activities. The player should be able to define the proper number of messages depending on the player's needs and interest in monitoring the game offline.

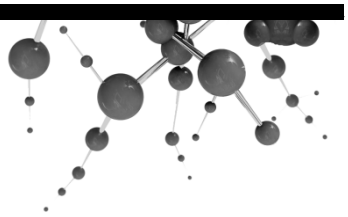
7 Conclusions

In this paper we have presented design challenges that game designers face when designing pervasive mobile games. We also present design implications how to overcome these challenges. These design implications were derived from a user study that was conducted with an experimental pervasive multiplayer mobile game. The game design utilized three types of context information in the game world: spatio-temporal, environmental, and proximity. Another important feature was an asynchronous slow update gameplay, which allowed the players to play the game whenever they wanted to and collaborate with other players without disturbing their gaming experience. The user study results indicated that when designing pervasive mobile games, the designer should pay attention to player's freedom to play the game. This means that the player should be able to adjust play sessions according to the available time they have. When the player is offline, communication outside the game world also becomes a crucial factor for convenient gameplay. Another important factor in pervasive games is the utilization of context information. The designers should ensure that both the players and the game system have mutual understanding of the current context. In addition, the game design should support that all players have equal chance to access relevant context information. By following these high level guidelines we believe that it is possible to design more enjoyable pervasive mobile games and introduce this new game genre to a wider player population.

Acknowledgements. IPerG (www.iperg.org) is an Integrated Project (FP6-004457) funded under the European Commission's IST Programme. We would like to thank all project members for their contribution to designing and implementing the game and the user study participants for their valuable comments.

References

1. Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., Benford, S., Capra, M., Hampshire, A.: Interweaving Mobile Games with Everyday Life. In: Proceedings of ACM SIGCHI, pp. 417–426 (2006)
2. Beyer, H., Holtzblatt, K.: Contextual Design: Defining Customer-Centered Systems, pp. 154–163. Morgan Kaufmann, San Francisco (1998)
3. Bogost, I.: Asynchronous Multiplay: Futures for Casual Multiplayer Experience. Other Players conference on Multiplayer Phenomena, The IT University of Copenhagen (2004), http://www.bogost.com/downloads/I.%20Bogost%20-%20Asynchronous%20Multiplay.pdf
4. Broll, W., Ohlenburg, J., Lindt, I., Herbst, I., Braun, A.: Meeting Technology Challenges of Pervasive Augmented Reality Games. In: Proceedings of 5th ACM SIGCOMM workshop on Network and system support for games (NetGames 2006) (2006)
5. Eriksson, D., Peitz, J., Björk, S.: Socially Adaptable Games. In: Proceedings of DIGRA 2005 Conference: Changing Views - Worlds in Play (2005)
6. Ghellal, S., Bullerdiek, S., Lindt, I., Pankoke-Babatz, U., Adams, M., Söderlund, T., Oppermann, L.: Design Guidelines for Crossmedia Game Production, Public IPerG Deliverable D8.1, http://www.pervasive-gaming.org/Deliverables/D8.1-Design-Guidelines-for-Crossmedia.pdf
7. Henricksen, K., Indulska, J.: Modelling and using imperfect context information. In: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops, pp. 33–37 (2004)
8. Koivisto, E.M.I.: Supporting Communities in Massively Multiplayer Online Role-Playing Games by Game Design. In: Proceedings of DIGRA Conference: Level Up (2003)
9. Lankoski, P., Heliö, S., Nummela, J., Lahti, J., Mäyrä, F., Ermi, L.: A case study in pervasive game design: the songs of north. In: Proceedings of the third Nordic conference on Human-Computer Interaction (NordiCHI 2004), pp. 413–416 (2004)
10. Lindley, C.A.: Game Space Design Foundations for Trans-Reality Games. In: Proceedings of the international conference on Advances in Computer Entertainment Technology (ACE 2005), pp. 397–404 (2005)
11. Linner, D., Kirsch, F., Radusch, I., Steglich, S.: Context-aware Multimedia Provisionning for Pervasive Games. In: Proceedings of the Seventh IEEE International Symposium on Multimedia (ISM 2005), pp. 60–68 (2005)
12. Magerkurth, C., Cheok, A.D., Mandryk, R.L., Nilsen, T.: Pervasive Games: Bringing Computer Entertainment Back to the Real World. ACM Computers in Entertainment 3(3) (2005)
13. Montola, M., Waern, A., Niewdorp, E.: Domain of Pervasive Gaming, Public IPerG Deliverable D5.3b (2006), http://iperg.sics.se/Deliverables/D5.3b-Domain-of-Pervasive-Gaming.pdf
14. Ohlenburg, J., Lindt, I., Pankoke-Babatz, U., Ghellal, S.: Report on the Crossmedia Game Epidemic Menace. ACM Computers in Entertainment 5(1) (2007)
15. Peitz, J., Saarenpää, H., Björk, S.: Insectopia – Exploring Pervasive Games through Technology already Pervasively Available. In: Proceedings of the international conference on Advances in Computer Entertainment Technology (ACE 2007), pp. 107–114 (2007)



Paper IV

Paavilainen, J., Korhonen, H. & Saarenpää, H. (2012). Comparing Two Playability Heuristic Sets with Expert Review Method: A Case Study of Mobile Game Evaluation. Chapter 2 in Lugmayr, A., Franssila, H., Näränen, P., Sotamaa, O., Vanhala, J., & Yu Z., (Eds.) *Media in the Ubiquitous Era: Ambient, Social, and Gaming Media*, 29-52. Hershey, PA, USA: IGI-Global. doi: 10.4018/978-1-60960-774-6.ch002

Media in the Ubiquitous Era:

Ambient, Social and Gaming Media

Artur Lugmayr
Tampere University of Technology, Finland

Helja Franssila
Hypermedia Laboratory, Finland

Pertti Näränen
Tampere University of Applied Sciences, Finland

Olli Sotamaa
Tampere University of Technology, Finland

Jukka Vanhala
Tampere University of Technology, Finland

Zhiwen Yu
Northwestern Polytechnical Unviersity, China

Information Science
REFERENCE

Senior Editorial Director: Kristin Klinger
Director of Book Publications: Julia Mosemann
Editorial Director: Lindsay Johnston
Acquisitions Editor: Erika Carter
Development Editor: Michael Killian
Production Editor: Sean Woznicki
Typesetters: Lisandro Gonzalez, Milan Vracarich
Print Coordinator: Jamie Snavely
Cover Design: Nick Newcomer

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Media in the ubiquitous era: ambient, social and gaming media / Artur Lugmayr ... [et al.], editors.
p. cm.
Includes bibliographical references and index.
Summary: “This book focuses on the definition of ambient and ubiquitous media from a cross-disciplinary viewpoint, covering the fields of commerce, science, research affecting citizens”--Provided by publisher.
ISBN 978-1-60960-774-6 (hbk.) -- ISBN 978-1-60960-775-3 (ebook) -- ISBN 978-1-60960-776-0 (print & perpetual access) 1. Ubiquitous computing. 2. Ambient intelligence. 3. Communication--Technological innovations. I. Lugmayr, Artur.
QA76.5915.M43 2012
004--dc23

2011031143

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 2
Comparing Two Playability Heuristic Sets with Expert Review Method:
A Case Study of Mobile Game Evaluation

Janne Paavilainen
University of Tampere, Finland

Hannu Korhonen
Nokia Research Center, Finland

Hannamari Saarenpää
University of Tampere, Finland

ABSTRACT

The expert review method is a widely adopted usability inspection method for evaluating productivity software. Recently, there has been increasing interest to apply this method for the evaluation of video games, as well. In order to use the method effectively, there need to be playability heuristics that take into account the characteristics of video games. There are several playability heuristic sets available, but they are substantially different, and they have not been compared to discover their strengths and weaknesses in game evaluations. In this chapter, we report on a study comparing two playability heuristic sets in evaluating the playability of a video game. The results indicate that the heuristics can assist inspectors in evaluating both the user interface and the gameplay aspects of the game. However, playability heuristics need to be developed further before they can be utilized by the practitioners. Especially, the clarity and comprehensibility of the heuristics need to be improved, and the optimal number of heuristics is still open.

DOI: 10.4018/978-1-60960-774-6.ch002

INTRODUCTION

Competition in the game industry is hard and the gaming experience has become a crucial factor in differentiating similar kinds of game titles. If a game is not enjoyable to play, players can easily switch to another game. Typically, gaming experience can be evaluated after there is a working prototype implemented and it is ready for beta testing. At this point, correcting any playability problems (e.g. UI navigation is complex, goals are not clear, or the challenge level or pace is set incorrectly) is often too expensive, or the project schedule does not allow any delays due to marketing reasons. As a result, there is a need for an evaluation method that can identify these playability problems before beta testing starts and thus provide time for corrections.

Productivity software has been evaluated for years with the expert review method to find usability problems in the design and implementation (Nielsen and Molich, 1990). In an expert review method, a small group of experts evaluate a product based on a set of heuristics. Heuristics are guidelines, rule of thumb statements, which reflect the desirable aspects of a given product. The method is cost-efficient and effective, and the design can be evaluated already in early project stages. A skillful and knowledgeable usability expert can identify usability problems as accurately as in user testing (Molich and Dumas, 2008). Evaluating games with this method is a tempting idea, but traditional usability heuristics cannot be applied directly (Federoff, 2002; Desurvire et al., 2004; Korhonen and Koivisto, 2006).

The design objectives between productivity software and games are different, and the evaluation methods need to recognize this divergence as well before they can be effectively applied to the domain of games. Pagulayan et al. (2008) describe these differences, and according to them, productivity software is a tool and the design intention is to make tasks easier, more efficient, less error-prone, and increase the quality of the results. Games, instead, are intended to be pleasurable to play and sufficiently challenging (Pagulayan et al., 2008). Because of these differences, a set of specifically designed heuristics are needed when video games are evaluated with the expert review method.

Playability has been studied very little by game researchers and HCI researchers. The research community is lacking a commonly agreed upon definition for playability, which would describe important issues influencing the game experience and guiding the research work. Egenfield-Nielsen et al. (2008) state that a game has good playability when it is easy to use, fun and challenging. Järvinen et al. (2002) have defined playability as an evaluation tool which consists of four components: 1) functional, 2) structural, 3) audiovisual, and 4) social playability. These components can be used to evaluate both the formal and the informal aspects of a game. Fabricatore et al. (2002) have defined playability in action games as the possibility of understanding and controlling gameplay. In addition, they state that poor playability cannot be balanced or replaced by non-functional aspects of the design. According to usability glossary[1], playability is affected by the quality of different aspects, including storyline, controls, pace, and usability. Along with the academia, the game industry has also approached the issue of playability from the practical perspective. For example, Games User Research at Microsoft Game Studios has published several empirical papers considering usability, playability and user experience in video games[2].

For our work we have defined playability as follows. Playability is related to intuitiveness, unobtrusiveness, fun, and challenge. In addition, it is a combination of user interface and the gameplay, i.e. game content aspects of the game. In multiplayer games, players' social interaction also affects playability. The user interface consists of game menus, controls and an interface through which a player interacts with game items, non-player characters (NPCs), and other players. A game has good playability when the user interface

is intuitive and unobtrusive, so that the player can concentrate on playing the game. Gameplay includes, for example, game mechanics, narrative, and goals that the player tries to achieve. Fun and challenge are created by the gameplay; a game has good playability especially when the gameplay is understandable, balanced, suitably difficult, and engaging.

Despite the lack of a commonly agreed upon definition, researchers have defined playability heuristic sets that could be used to evaluate video games and their playability. However, the development work is still ongoing and there is very little knowledge about the usefulness and clarity of these heuristic sets. In addition, there are no previously published studies that would use these heuristic sets to evaluate a video game and compare the results.

In this chapter, we report an experiment in which two playability heuristic sets are used in a video game evaluation to discover their weaknesses and strengths in identifying playability problems, as well as whether they are helpful to inspectors in conducting the evaluation. The results indicate that heuristic sets should be improved before they are usable for the practitioners.

The rest of this chapter is structured as follows. First, we review relevant related work regarding the expert review method and introduce playability heuristics that have been developed. Next, we describe an experiment we arranged to compare two playability heuristic sets in game evaluation and report the results of the experiment followed by discussion and conclusions.

RELATED WORK

In this section, we present the expert review method and look at playability heuristics that have been developed for evaluating video games.

The Expert Review Method

Inspection methods are well known and widely used to evaluate the usability of products. According to Nielsen (2005a) heuristic evaluation is the most popular usability inspection method. The popularity of the method is due to its cost-effectiveness and ease of implementation in discovering usability problems even by novice inspectors. The method was developed by Nielsen and Molich (1990) and it is also known as the expert review method, since the inspectors' experience and knowledge affect the results of the evaluation (Jacobsen et al., 1998). The first version of usability heuristics was published together with the method, but the revised and currently used version of the heuristics was published in 1994 (Nielsen, 1994a).

When conducting an expert review, inspectors go through the evaluated software individually and write down usability problems that they notice. The problems are assigned with the appropriate heuristic, and later the findings from all inspectors are synthesized into a single usability report with suggestions on how to fix the problems. The heuristics act as a guideline for the inspectors to focus on typical usability issues that cause problems. Moreover, the problems found are also rated with a severity rating (e.g. three-scale rating minor, major, and critical) to emphasize its severity for the usability of the product. Nielsen (2005b) has suggested that three to five expert inspectors (preferably with domain expertise) are enough, since adding more inspectors would not increase the amount of problems found significantly.

Nielsen's heuristics were made for evaluating productive software interface design. However, several researchers have extended these heuristics or developed new ones for different application domains. Ling and Salvendy (2005) present a summary of some of these studies. The summary contains domains such as websites, e-learning systems, groupware, notification systems, and games.

The applicability of traditional usability heuristics in game evaluations has been questioned by game researchers (Federoff, 2002; Desurvire et al., 2004; Korhonen and Koivisto, 2006). The most important reason for this is that usability heuristics concentrate primarily on the user interface and disregard other aspects of a product. For example, in video games it is equally important to evaluate the gameplay as well. In their study, Johnson and Wiles show how games contravene the traditional usability heuristics to achieve a good game experience (Johnson and Wiles, 2003). Hence, game researchers have started to develop heuristics which would include both usability and gameplay issues, to assist game developers in discovering playability problems in the game design.

Development of Playability Heuristics

In the early 1980s, Malone (1982) studied video games and what makes the user interface enjoyable. He identified three principles (challenge, fantasy, and curiosity) that are needed for designing enjoyable user interfaces. Malone also calls these principles heuristics in a design framework. Although the list is very limited and it concentrates only on high level issues in games, it highlights the importance of the game content in the evaluation.

Clanton argued that human-computer interaction in games can be divided into three levels: game interface, game mechanics and game play (Clanton, 1998). Furthermore, Clanton describes 15 principles which can be used to gain and keep the interest of a player. Although these principles are not called heuristics, they can be understood as such.

Federoff (2002) defined the first playability heuristics that are similar to usability heuristics. These heuristics were a result of a case study in a game company, but they lack validation, or at least such results have not been published. Fabricatore et al. (2002) have studied players and their preferences that will affect the playability of action video games. Even though these are not described as heuristics, they could be converted into heuristics to evaluate games belonging to this specific genre.

Desurvire et al. (2004) published Heuristic Evaluation for Playability (HEP) in 2004. The heuristics were validated in a study in which heuristic evaluation was compared to user testing. The results indicated that the HEP heuristics were very good at identifying playability problems in a game prototype.

Korhonen and Koivisto (2006) have published a playability heuristic set for mobile games. However, the heuristics are applicable for evaluating games in other platforms as well because of their modular structure. The playability heuristic set can be extended or limited based on the needs of the evaluation. Later on, the core heuristics were extended with multiplayer heuristics (Korhonen and Koivisto, 2007). In addition, the number of heuristics is smaller than in two previous sets from Federoff (2002) and Desurvire et al. (2004). The heuristics of Korhonen and Koivisto have been validated in several game evaluations.

Schaffer (2007) presented a white paper introducing heuristics for usability in games. According to Schaffer, earlier heuristics lacked concrete examples, making them less clear for practitioners. Schaffer's heuristics are based on literature and on his own expertise from the field of human-computer interaction.

Pinelle et al. published ten game usability heuristics that are based on game reviews (Pinelle et al. 2008a) and they have been validated in a preliminary study. These heuristics are used to evaluate game usability and there are no heuristics concerning gameplay issues. Later Pinelle et al. (2009) also published ten additional heuristics focusing on multiplayer usability.

Desurvire and Wiberg (2009) have presented PLAY heuristics, which are based on the earlier HEP heuristics (Desurvire et al., 2004). The PLAY heuristics feature 19 top level headings, each containing one to six heuristics (50 heuristics in

total). The PLAY heuristics are aimed toward action adventure, first person shooter and real-time strategy games.

More recently, Köffel et al. (2010) have presented a synthesis of earlier heuristics. The authors handpicked 29 heuristics from the earlier models and added ten more heuristics focusing on advanced electronic tabletop games.

There are also other guidelines that are targeted for game developers in order to make games more engaging and usable for players (e.g. Falstein and Barwood, 2001; Snow, 2007). In addition to articles and websites, edited books have also been published recently on the topic (e.g. Isbister and Schaffer, 2008; Bernhaupt, 2010). These cover various methods for evaluating usability, playability and user experience in general.

Based on the literature review, expert review could be an appropriate method for evaluating the playability of video games, but there should be specific playability heuristics accompanying the method. Several researchers have started to develop these heuristics, and currently, there are multiple heuristic sets available. However, the work is still ongoing and the heuristic sets are quite different, even though there are some common issues included. This raises the question of which heuristic set should be used in a game evaluation, and if one heuristic set is easier to use than another one from the inspectors' point of view.

In our work, we aim to achieve some clarity about the different playability heuristic sets and their usefulness in game evaluations. We compare two playability heuristic sets in a game evaluation to see what their strengths and weaknesses are and how inspectors perceive the heuristic sets.

Validation of Domain Specific Heuristics

Traditional usability heuristics (Nielsen, 1994a) are widely used in usability evaluations. One limitation of these heuristics is that they are advisedly general and they do not cover specific characteristics of systems (Ling and Salvendy, 2005: Sim et al., 2009). Nielsen has noted that there could be domain-specific heuristics for a specific class of products as a supplement to the general heuristics.

Several researchers have developed heuristics for different domains. Baker et al. (2001) have developed heuristics to identify usability problems in real-time collaboration within a shared visual workspace. In their validation study, two groups of inspectors evaluated two groupware applications by using these heuristics. The evaluation results were compared to previously published studies by replicating Nielsen and Molich's analysis methodology (Nielsen, 1990; Nielsen, 1992; Nielsen, 1994a).

Berry (2003) has developed heuristics for notification systems and compared them to traditional usability heuristics. In the study, inspectors were divided into two groups and they evaluated three versions of the system's user interface. The results indicate that both heuristic sets performed quite similarly in identifying usability problems.

Mankoff et al. (2003) compared the performance of heuristics developed for evaluation of ambient displays to traditional usability heuristics. The results indicate that with the help of modified heuristics inspectors were able to identify more usability problems than inspectors using traditional usability heuristics. However, the best result was achieved by combining both sets of heuristics.

Bertini et al. (2009) have also developed mobile usability heuristics that not only take into account applications, but also the device and context in which it is used. This heuristic set was also evaluated against the traditional usability heuristics in the evaluation of two mobile applications.

Previously presented studies compared domain-specific heuristics to traditional usability heuristics. Although this kind of comparison is useful in determining the usefulness of a new heuristic set, it still lacks a critical analysis of the heuristics to determine their usefulness in that specific domain. One of the problems related

to such comparison studies is that there are not many domains that have multiple domain-specific heuristic sets. Ling and Salvendy (2005) have presented a summary of some studies in which domain-specific heuristic sets have been developed. In their review, each domain contained only one heuristic set.

Zuk et al. (2006) used three sets of domain-specific heuristics that are targeted to information visualization systems. The goal of the study was not to compare heuristic sets as such, but to identify a common set of heuristics derived from these three heuristic sets to find common visualization problems.

Video games are one of the few application domains that have multiple heuristic sets developed for them (e.g. Federoff, 2002; Desurvire et al., 2004; Korhonen and Koivisto, 2006; Pinelle, 2008a). However, studies empirically evaluating different heuristic sets in this domain and comparing the applicability of the heuristic sets have not been published.

THE EXPERIMENT

We arranged an evaluation session with eight persons who are working in the game industry or in the academia as game researchers to explore how the expert review method and two playability heuristic sets operate in a game evaluation. First, the participants were briefly introduced to the expert review method and the heuristics that are commonly used in productivity software evaluations to give an idea of how usability specialists usually conduct evaluations. Three participants had previous experience in conducting an expert review of a product.

The participants were divided into 4 teams (two persons in each team) forming two groups based on the playability heuristic sets that were given to them. In the evaluation session, the teams played a game for one hour. The inspectors observed the game and wrote down short descriptions of possible playability problems they encountered in the game. After that, the teams went through their own playability problems and assigned violated playability heuristics to these problems. Finally, the observations were talked through with other teams and the participants discussed playability problems, the evaluation method, and the playability heuristics they used. The results section describes the main observations from the discussion, which was recorded with a video camera.

The game evaluated was EA Mobile's The Simpsons: Minutes to Meltdown[3]. We selected this game for the evaluation because it was short enough to be evaluated in a single day workshop as it can be completed in less than 30 minutes in real time. The game did not receive favorable reviews (e.g. Buchanan 2007; Dredge, 2007; PurestProdigy, 2008), which made us to believe there were many playability problems to be found. We also wanted to use a mobile game, because it would be easy to obtain for every inspector and the evaluation session would be easy to arrange.

In this game, the player controls Homer who has 30 minutes to save Springfield from a nuclear disaster. The game features slightly tilted pseudo-3D top-down perspective and Homer's movement is controlled with a rocker key, or keys 2, 4, 6 and 8 (up, left, right, down respectively) on the keypad. Left soft key brings up the pause menu and right soft key shows the timer. Context sensitive action is executed by pressing the rocker key or key 5 on the keypad.

The game features three levels and locations. The game starts at the Simpson's apartment where Homer must find his car keys so he can drive to the power plant. Homer must interact with various characters and avoid furious citizens to complete the first level. The second level is the Springfield city centre. Homer has crashed his car and must continue on foot. There are various obstacles and hazards on the way as he tries to reach the plant. The last level is the power plant, where Homer must find the override valve to cancel the meltdown.

In this level, depending on the mobile phone, the player also controls Bart in certain key locations.

Playability Heuristics

As there are multiple heuristic sets available, it is important to choose heuristic sets that can be compared. As described in the chapter on related work, some heuristics are proposals which have not been validated, others are targeted to a specific game genre or they do not consider all aspects of playability. For this study, we selected playability heuristic sets from Desurvire et al. (2004) and Korhonen and Koivisto (2006) because they resembled each other, they were probably the most advanced at the time of the study and they have both been validated by their authors. Although there were other heuristics available, they were either not validated by their authors (e.g. Federoff, 2002; Schaffer, 2007) or they only focused on usability issues in games (e.g. Pinelle et al., 2008a).

The heuristic sets selected are based on literature reviews and the initial heuristics were reviewed by game researchers and game designers. The playability heuristics were developed further in game development projects and they were both validated by their authors. Although the sets have some similarities in their content, there are major differences in how the heuristics are organized and described.

Heuristic Evaluation of Playability (HEP)

This playability heuristic set contains 43 heuristics and the authors have defined four categories for organizing them (Desurvire et al., 2004; see Appendix 1). *Game Play* is related to challenges and problems that the player must face to win a game. *Game Story* includes heuristics for story and character development. *Game Mechanics* involves the structure which defines how the game units interact with the environment. *Game Usability* addresses the interface and the controls the player utilizes when interacting with the game. Most heuristics are presented as one sentence descriptions and they have been validated in a user study. Teams that used this playability heuristic set during the evaluation are referred to as Violet 1 and Violet 2 in the results section.

Playability Heuristics for Mobile Games

This playability heuristic set contains 29 heuristics which have been organized into three modules (Korhonen and Koivisto, 2006; see Appendix 2). Each module can be included or excluded depending on the needs of the evaluation. Two core modules, *Gameplay* and *Game Usability*, are common to all games. The *Mobility* module contains heuristics that are specific to mobile games. Each heuristic is described in detail on a separate document including examples of use (Koivisto and Korhonen, 2006). The heuristics were validated in several mobile game evaluations conducted by playability experts. Teams that used this playability heuristic set during the evaluation are referred to as Orange 1 and Orange 2 in the results section.

RESULTS

In this section, we present the main results of the study, which are based on the comments from the group interview as well as the analysis of the data collected from the evaluation reports.

Heuristics Provide Guidance

The participants commented that the expert review method seemed to be an appropriate method for evaluating video games, because it helped them to focus on the different aspects of the game during the evaluation. One game industry participant stated that they use similar kind of evaluation approach on a weekly basis to manage game production processes.

Participants commented that heuristics could also be useful in the design and implementation phases in game development to identify possible playability problems that might exist in the design. The participants would not use heuristics in the very early phases of game development, as they considered that the heuristics might restrict creativity at that point. One double expert participant also stated that she would not use heuristics in the very late phases either, as actual playtesting would be more beneficial. According to her, heuristics are good for finding basic problems and playtesting is used to refine the details.

Defining a Proper Abstraction Level

Although the participants appreciated the efficiency of the expert review method, they stated that there are certain challenges when the method is applied to game evaluations. Their biggest concern was related to the heuristics and their descriptions. The variety of video games is enormous and defining playability heuristics that are suitable for evaluating all kinds of games can be a challenge.

"It is a laborious and challenging task to define heuristics that can capture those aspects that are considered to be essential from the point of view of game experience. In addition, game environments are changing constantly as they adopt new kinds of technical enablers", Violet 1 inspector.

Therefore, it is important that the playability heuristics are on the right abstraction level. Too specific heuristics restrict their applicability to a large number of games, but in contrast, heuristics that are on a very general level lose their power to guide and assist inspectors during the evaluation. The participants stated that both heuristic sets had problems in this respect.

Playability heuristics defined by Desurvire et al. had both detailed heuristics and very broad heuristics, which were difficult to use during the evaluation. For example, there is the Game Play heuristic number 10 *("The game is fun for the Player first, the designer second and the computer third. That is, if the non-expert player's experience isn't put first, excellent game mechanics and graphics programming triumphs are meaningless.")* This heuristic was considered to be very difficult to apply during an evaluation.

Playability heuristics defined by Korhonen and Koivisto also had some heuristics which were considered to be quite specific, and they could be combined to provide a more concise list. For example, heuristics GP9 (*"The players can express themselves"*) and GP10 (*"The game supports different playing styles"*) describe similar kinds of issues on heading level, which are related to the player's behavior and playing style in the game world.

Usability and Gameplay Problems

The participants suggested that it is generally easier to find usability problems than it is to find gameplay problems from games. However, in this case, the participants found more gameplay related problems. One suggested reason for this was that the inspectors were experienced gamers and might thus be blind to some usability problems that would annoy novice gamers. On the other hand, in this case the user interface was generally liked as it was considered to reflect the industry standards well. From this perspective, it might be easier to spot usability problems, which are few but obvious, than playability problems, which are more numerous but also harder to point out. Korhonen and Koivisto (2006) reported similar findings in their earlier study when they developed their heuristics.

One participant questioned the difference between usability and gameplay problems. Sometimes, it was hard to decide if the problem found was related to usability or gameplay. *"There were a lot of things which were on the borderline of being in the game's interaction structure or in the user*

interface structure. In the end, it was not possible to separate them clearly", Orange 2 inspector.

In addition, it was discussed that it is not always obvious whether gameplay problems are actually problems or just hard challenges. For example, if a player has to go through the same part of the game over and over again, but is highly immersed and enjoys the game, calling out a gameplay problem might not be necessary. In our case, it was noted by one inspector that the game evaluated featured so many problems that it was not possible to get immersed.

Evaluation Process

The participants commented that the evaluation task influenced their gaming experience, and for that reason, playing the game was different than what it would be normally. The objective of the game design is to immerse players on different levels (Ermi and Mäyrä, 2005). The evaluation task, however, prevents immersion because the inspectors need to be alert all the time and inspect the game for problems in playability. In addition, the inspectors found it difficult to play the game like any other player would, and for that reason, the evaluation session cannot be considered equal to a normal play session.

"There are two dimensions that make the evaluation difficult. First, you should be able to describe the problem that you have identified and it affects your gaming experience negatively. On the other hand, you should play the game as players would play and get a positive gaming experience", Violet 1 inspector.

Another issue the participants pointed out was that it is very important for the inspectors to familiarize themselves with the heuristics beforehand. In our study, playability heuristics sets contained 43 or 29 heuristics. When considering Miller's golden rule of 7±2 (Miller, 1956), the number of heuristics might have been overwhelming and there was too much information about the heuristics to keep in mind. During the evaluation, it was time-consuming to browse through the whole list and find a proper heuristic for each playability problem. Due to time constraints, the participants did not study the heuristics beforehand, but there was a playability expert present in case they had any questions concerning the heuristics.

The large amount of heuristics brought up an idea in the end discussion that inspectors could use the heuristics in a more systematic fashion. First, the game would be played for some time, and then the inspector would go through the heuristics in a checklist manner. This approach has also been suggested earlier by Nielsen (1994b). However, it was noted that in this case the tool is using the inspector and not vice versa and therefore there is a possibility that the inspector does not recognize possible problems that are outside of the scope of the heuristics. It was also considered that the heuristics in general are problem-oriented and do not support positive findings very well. It was suggested that instead of formal statements, the heuristics could be in the form of questions, which might intuitively help to discover positive features from the game. One participant stated that due to his expertise in game development, he was able to find positive features easily.

Finding positive features was considered important. One inspector stated that, from a psychological perspective, no one likes to read an evaluation report which is full of negativity. Especially if a report is delivered to a person who has not seen the game, it might give a false impression of it. Reporting positive features of the game also enhances the possibility that those features are left intact and are not accidentally removed, changed or "fixed" by the designers (Nielsen 1994a).

Revision for Playability Heuristic Sets

The participants found several issues troublesome with the playability heuristics defined by Desurvire et al. These issues made utilization of the heuristics difficult during the evaluation. There is a total of 43 heuristics in the set and the participants thought this is too much. The heuristics are organized into four categories, but the participants did not find them helpful because some heuristics were in a different category from what they expected. For example, some Game Story heuristics were located in the Game Play category and vice versa. The Violet team inspectors pointed out that Game Play heuristic number 8 (*"Players discover the story as part of game play"*) would belong to the Game Story category rather than the Game Play category, and that Game Story heuristic number 6 (*"Player experiences fairness of outcomes"*) sounds more like a heuristic belonging to the Game Play category. There were also some overlapping heuristics in the set.

Another problem that the participants noticed was the descriptions of the heuristics, as they were presumably influenced by the game that was used as a basis during the development work. Some heuristics were seen to be too specific to apply in practice. In addition, the descriptions were not consistent in terms of wording and the level of generalization. Some heuristics clearly set requirements for the game design and state explicitly how the game should be designed. An example of this kind of a heuristic is Game Play heuristic number 3 (*"Provide clear goals, present overriding goal clearly as well as short-term goals throughout the play"*), whereas some heuristics are more like recommendations for designers. For example, Game Play heuristic number 5 (*"The game is enjoyable to replay"*) is a too general and subjective issue to evaluate with the expert-based method. There were also some heuristics which were difficult to understand and apply during the evaluation. The participants pointed out Game Play heuristic number 10 *("The game is fun for the Player first, the designer second and the computer third. That is, if the non-expert player's experience isn't put first, excellent game mechanics and graphics programming triumphs are meaningless"*) to be an example of such a heuristic. Finally, the participants commented that the current writing style and format makes understanding the heuristics more difficult because they are not consistent and are missing either a heading or a description.

Playability heuristics developed by Korhonen and Koivisto were not optimal either. Even though each heuristic clearly had the heading and the description, they were presented in two documents which made using them difficult. The first document described the heuristics on a heading level, in a similar fashion to the other heuristic set. There was a separate document available that contained the descriptions and practical examples (Koivisto and Korhonen, 2006). Some descriptions were also long, and reading the entire description and examples was time-consuming. The participants commented that this playability heuristic set was in a better shape and the wording of the heuristics was more consistent and on a more generic level than in the other heuristic set. However, there were still some heuristics such as GP8 (*"There are no repetitive or boring tasks"*) and GP11 (*"The game does not stagnate"*) that sounded similar on the heading level and they could possibly be combined.

One suggestion was that heuristics could be organized on different levels inside one category. For example, high level usability problems would consist of more abstract heuristics which are applicable to a large number of games. Low level usability heuristics would be more focused on certain game genres as it has been noted that different genres have different problems (Pinelle, 2008b).

Table 1. Playability problems concerning different heuristic categories

		Orange				Violet			
	Teams	**O1**	**O2**	**Total**	**%**	**V1**	**V2**	**Total**	**%**
Categories	Game Usability	3	4	7	28%	6	3	9	20%
	Gameplay	8	5	13	52%	3	11	14	32%
	Mobility	1	0	1	4%	-	-	-	-
	Game Story	-	-	-	-	1	2	3	7%
	Game Mechanics	-	-	-	-	5	0	5	11%
	Unassigned	1	3	4	16%	5	8	13	30%
	Total	13	12	25	100%	20	24	44	100%

Evaluation Statistics

Surprisingly, there was very little consistency in reporting playability problems between the four teams. Only a few playability problems were identified by more than one team. Even teams with the same heuristic lists assigned different heuristics to a playability problem. The teams reported 69 playability problems in total. 13 playability problems were reported by two or more teams and 52 playability problems were uniquely reported by a single team. In addition, there were 13 duplicate playability problems (i.e. reported multiple times by a single team), but these problems have been excluded from the analysis. There was a difference between groups in how many playability problems they reported. Teams Orange 1 (O1) and Orange 2 (O2) identified 13 and 12 playability problems respectively. Teams Violet 1 (V1) and Violet 2 (V2) identified a substantially larger number of playability problems, 20 and 24 playability problems respectively.

Most problems reported by both teams were related to gameplay issues. Teams O1 and O2 reported more than a half of the problems belonging to this category. The second most common problem category was game usability. Playability problem distribution in the heuristic categories is illustrated in Table 1. It should be noted that the heuristic categories are not comparable because they contain different heuristics. In addition, some categories exist only in one playability heuristic set and those categories are left empty on the table in the other set.

Some user interface problems were due to the mobile phones the participants used. The game looked and sounded different on their devices, and there were some minor changes in the game content because of the smaller screen resolution and the memory capacity of the device.

The teams seemed to have difficulties in assigning violated heuristics to the identified playability problems, and the participants commented that they could not always find a proper playability heuristic from the set. Especially for teams V1 and V2, assigning a violated playability heuristic was difficult, and they left 30% of the playability problems open (Table 1). Teams O1 and O2 were able to do it more accurately, and they left only 16% of reported playability problems open.

Usually the teams assigned only one violated heuristic per problem, but there were a few cases when they assigned several heuristics (Table 2). The teams reported nine playability problems to which they assigned several heuristics from the same category that the problem violated. Three of them were related to Game Usability and the rest were Gameplay problems. There were also three playability problems to which the teams assigned playability problems from different categories. These problems were combinations of Gameplay, Game Usability, and Game Story related issues.

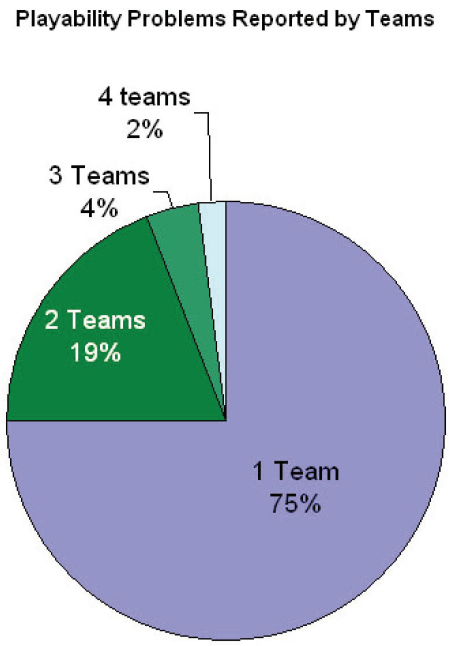
Table 2. Assigning heuristics to playability problems

		Orange Teams		Violet Teams	
		Count	%	Count	%
Heuristics	Single	12	48%	28	64%
	Many Same Category	8	32%	1	2%
	Many Different Category	1	4%	2	5%
	Unassigned	4	16%	13	30%
	Total	25	100%	44	1005

Finding the same playability problems seemed to be difficult, and the majority of the playability problems (75%) are reported only by a single team (Figure 1). However, there was one playability problem which all teams reported. The playability problem concerns player progression in the game. If Homer dies in the game, the player has to start all over from the beginning. Teams were also consistent when assigning the violated heuristic for this problem. Teams O1 and O2 assigned the gameplay heuristic GP14 (*"The player does not lose any hard-won possessions"*). In addition, Team O1 marked that the problem violated Gameplay heuristic GP8 (*"There are no repetitive or boring tasks"*). Teams V1 and V2 also had a consensus on the violated heuristic. They assigned Game Story heuristic GS6 (*"Player experiences fairness of outcomes"*) to describe the problem. In addition, team V1 assigned Game Play heuristic GP5 (*"The game is enjoyable to replay"*).

There were two problems that were identified by three teams. The first problem concerned navigation in the game world, due to the fact that the player gets lost very easily on the second level. The second problem was related to the game menu design. Even though the three teams identified the same problem, each team assigned a different heuristic to describe the problem or left the problem open. For the playability problems identified by two teams, there was hardly any consistency in the heuristics assigned. One explanation for different evaluation results between

Figure 1. Playability problems reported by teams



teams O1 and O2 might be that team O1 reported very specific playability problems such as *"catching the pig is hard and it is not clear how it should be done"*, whereas team O2 reported more general level problems like *"the game is too linear and prone to stagnate"* or *"game features boring repetition without optional ways to advance"*.

Similarly, teams V1 and V2 used quite different heuristics to describe the playability problems. Team V2 did not assign any playability problems to the Mechanics category, even though team V1

used Mechanics quite extensively. They found three playability problems that violated heuristic ME1 *("game should react in a consistent, challenging, and exciting way to the player's actions (e.g., appropriate music with the action)"*). In addition, they assigned two other heuristics from the Mechanics category to describe identified problems. Correspondingly, team V2 concluded that five playability problems violated Game Play heuristic GP2 (*"provide consistency between the game elements and the overarching setting and story to suspend disbelief"*), while team V1 thought that none of their playability problems violated this heuristic.

Both teams reported playability problems with different abstraction levels. Team V1 identified both specific and general level problems, whereas team V2 concentrated on criticizing the illogical gameplay. Examples of such playability problems in the gameplay were *"Barney opens up a gate when you bring him coffee"* and *"The player can only go through certain bushes"*.

DISCUSSION

The inspectors' comments indicate that the expert review method is applicable to game evaluations. This supports earlier claims made by several researchers (e.g. Desurvire et al., 2004; Korhonen and Koivisto, 2006; Laitinen 2006; Pinelle et al., 2008a). The inspectors liked the method as it is not too time-consuming or laborious to execute. They thought that the method could also be used at earlier development phases, when there are only design sketches or low fidelity prototypes available. Playability heuristics, however, need to be developed further before the method can be widely adopted by the practitioners. Playability heuristics should be presented in a similar manner to how Nielsen (1994a) has presented traditional playability heuristics. Ling and Salvendy (2005) have also concluded that domain-specific heuristic sets should be structured and they should not contain too many heuristics.

In this study, we used playability heuristic sets developed by Desurvire et al. (2004) and by Korhonen and Koivisto (2006). The study revealed that both heuristic sets need to be improved in order for them to be usable and easily understandable. The inspectors considered that there were too many heuristics in the set developed by Desurvire et al (2004). In addition, their organization into categories, as well as their descriptions, need to be developed further as they were inconsistent and overlapping. This was visible in the evaluation data, as the teams who used this heuristic set did not assign any violated heuristic to 30% of the identified playability problems.

The playability heuristic set developed by Korhonen and Koivisto (2006) was more consistent in wording and organization, but the inspectors thought that the heuristics should be accompanied by short and compact descriptions since the descriptions were presented in a separate document.

In the study design, it is important to think about the hardware that will be used, since it can have remarkable influence on what kinds of playability problems are reported. Especially mobile phones can be very different in their technical capabilities and there are many device generations on the market. We did not anticipate that the game would vary so much on different devices. In our study, the inspectors used their personal mobile phones in the evaluation and therefore, we did not have sufficient control over the hardware. Some teams reported playability problems which were somewhat specific for the device they used. These problems were related to the audio and the amount of content on the screen. Gray and Salzman (1998) call this as an internal validity problem.

In game evaluations, the inspectors seem to face similar challenges in identifying the same playability problems. This result is consistent with comparison studies conducted with productivity software. However, the results of this study are slightly better than those reported by Molich and

Dumas (2008). The majority of the playability problems (75%) were reported only by a single team. However, one playability problem was commonly reported by all four teams, and the violated heuristic was assigned consistently within the teams, as well. Furthermore, there were 12 playability problems which were reported by at least two teams. It is an interesting question for future work why the inspectors do not identify the same problems in the game. Unlike productivity software, video games in general are quite linear at the beginning and the players are guided through first missions or levels by the game design (Adams and Rollings, 2007). Therefore, the inspectors should have gone through the same aspects of the game and presumably identified the same problems. The problem that all teams identified in this study was critical for the game experience, and this is probably the reason why it was reported. Further research is required to understand why the teams did not find the same playability problems.

There are several possible reasons for the inconsistency of the problems reported. One obvious explanation is the inspector effect (Jacobsen et al., 1998) and its influence on the results. It has been concluded in many previous evaluation studies of productivity software that evaluation results differ quite a lot because of this factor (e.g. Jacobsen et al., 1998; Molich and Dumas, 2008). In this study, our inspectors had different backgrounds, game design, and evaluation experience. Although we tried to balance teams in their evaluation experience and game design experience, it did not seem to be enough.

Another possible explanation for the inconsistency might be the heuristic sets that the inspectors used in the study. The purpose of the heuristic sets is to guide the evaluation and remind the inspectors to pay attention to important aspects of playability. The results indicate that using the heuristic sets was not straightforward and the inspectors had some problems with them, which might also explain the difference in reported playability problems. However, one interesting observation from the data is that most of the playability problems that were reported by two or three teams included teams from both groups. There were only a few problems which were reported only by one group. Unfortunately, there is not sufficient data from this study to make any deeper analysis on how a playability heuristic set influences finding playability problems in a video game.

Third possible explanation for this inconsistency might be that the inspectors had a different baseline for reporting. Some teams mainly reported general problems, focusing on certain aspects of the game, while others reported very specific problems. In the Violet group, team V2 did not report any playability problems which would violate heuristics from the Mechanics category, whereas team V1 assigned five problems to this category. Correspondingly, team V1 assigned five playability problems to one Game Play heuristic which was not used by team V2 at all. Otherwise, the teams reported problems that violated a large number of playability heuristics from different categories. This difference is probably due to the evaluation experience that the teams had. In addition, we probably did not instruct the teams clearly enough on what kinds of issues they should pay attention to and how to report those findings. In future studies, there should be greater emphasis on the instructions for creating a problem reporting baseline as equal as possible.

One characteristic of the game evaluations is to think about the origin of the playability problem, and whether the problem is in the user interface or in the game content. This problem does not usually exist in productivity software evaluations, as the evaluation concerns only user interface aspects of the product. Evaluating the content and the user interface together has been studied on other domains (e.g Galagher et al., 2001). In our study, the inspectors identified 12 playability problems to which they assigned multiple heuristics, and in three cases they were from different categories. We do not know for sure why the inspectors did

it this way. They possibly did not have time to analyze the problem thoroughly to find the origin of the problem.

It was also noted that the inspectors should always be aware of the creative vision that the designers have and what is used as a design principle when designing a game. Typically, it also guides the experience that the designers want to create for the players (Pagulayan and Steury, 2004). The evaluation should always be relative to this vision, because otherwise the inspectors might focus the evaluation incorrectly and point out issues which are contradicting to the vision. This further emphasizes the need for the designers and the inspectors to work in close cooperation.

In the future, we are planning to continue these comparison studies to find out the optimal set of playability heuristics. The shortcoming of this study was that we could not compare which playability heuristics are used to describe the same playability problems because there was too little data for this. In the next study, we should also eliminate internal validity errors, which were related to the inspectors' experience in using the evaluation method, their familiarity with the playability heuristics, and the devices that the inspectors used in the evaluation. In addition, there is a need to think about a new presentation format for the heuristics, which would better support the evaluation. In the discussion, it became obvious that presenting heuristics as a list is not easily utilized during the evaluation. The heuristics could be improved by using keywords, color coding for the categories and presenting them in a compact format, such as cards, for example.

CONCLUSION

In this chapter, we have explored two different playability heuristic sets to discover their strengths and weaknesses when they are used to evaluate a mobile game using the expert review method. This kind of a comparison study has not been reported previously, although there are several playability heuristic sets available. The results indicate that both heuristic sets should be improved as there were problems in clarity and comprehensibility. This study is the first attempt to develop playability heuristics that would help inspectors to conduct game evaluations, and to provide precise and relevant evaluation results when evaluating video games with an analytical evaluation method.

ACKNOWLEDGMENT

The GameSpace project was funded by Tekes (Finnish Funding Agency for Technology and Innovation), Veikkaus, TeliaSonera Finland, Nokia, Sulake Corporation, and Digital Chocolate. We thank all project members and the inspectors in this study.

REFERENCES

Adams, E., & Rollings, A. (2007). *Game Design and Development: Fundamentals of Game Design*. Upper Saddle River, NJ: Prentice Hall.

Baker, K., Greenberg, S., & Gutwin, C. (2001). Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. In *Proceedings of the 8th IFIP International Conference on Engineering for Human-Computer Interaction* (pp. 123-140). Springer-Verlag. Retrieved September 15, 2010, from http://portal.acm.org/ citation.cfm?id=645350.650731

Bernhaupt, R. (Ed.). (2010). *Evaluating User Experience in Games: Concepts and Methods*. London, UK: Springer-Verlag.

Berry, B. (2003). *Adapting Heuristics for Notification Systems*. Paper presented at the ACM SE Conference.

Bertini, E., Catarci, T., Dix, A., Gabrielli, S., Kimani, S., & Santucci, G. (2009). Appropriating Heuristic Evaluation for Mobile Computing. *International Journal of Mobile Human Computer Interaction*, *1*(1), 20–41. doi:10.4018/jmhci.2009010102

Buchanan, L. (2007). The Simpsons: Minutes to Meltdown Review. *IGN game review*. Retrieved September 10, 2010, from http://wireless.ign.com/articles /807/807589p1.html

Clanton, C. (1998). An interpreted demonstration of computer game design. In *CHI 98 conference summary on Human factors in computing systems* (pp. 1-2). Los Angeles, CA: ACM.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1509-1512). Vienna, Austria: ACM.

Desurvire, H., & Wiberg, C. (2009). Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration. In A. Ozok & P. Zaphiris (Eds.), *Online Communities and Social Computing*, Lecture Notes in Computer Science (Vol. 5621, pp. 557-566). Springer Berlin / Heidelberg. Retrieved September 15, 2010, from http://dx.doi.org/10.1007/978-3-642-02774-1_60

Dredge, S. (2007). The Simpsons: Minutes to Meltdown. PocketGamer game review. Retrieved September 10, 2010, from http://www.pocketgamer.co.uk/r/Mobile/The+Simpsons:+Minutes+To+Meltdown/review.asp?c=3683

Egenfield-Nielsen, S., Smith, J. H., & Tosca, S. P. (2008). *Understanding Video Games: The Essential Introduction*. London, UK: Routledge.

Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. In *DIGRA 2005 Conference: Changing Views – Worlds in Play*. Retrieved September 15, 2010, from http://citeseerx.ist.psu.edu /viewdoc/summary?

Fabricatore, C., Nussbaum, M., & Rosas, R. (2002). Playability in Action Videogames: A Qualitative Design Model. *Human-Computer Interaction*, *17*(4), 311–368. doi:10.1207/S15327051HCI1704_1

Falstein, N., & Barwood, H. The 400 Project. Retrieved September 10, 2010, from http://theinspiracy.com/ 400_project.htm.

Federoff, M. A. (2002). *Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games*. Master's Thesis, Department of Telecommunications, Indiana University, Indiana.

Galagher, K., Foster, D., & Parsons, J. (2001). The Medium Is Not the Message: Advertising Effectiveness and Content Evaluation in Print and on the Web. *Journal of Advertising Research*, *41*(4), 57–70.

Gray, W. D., & Salzman, M. C. (1998). Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. *Human-Computer Interaction*, *13*(3), 203–261. doi:10.1207/s15327051hci1303_2

Isbister, K., & Schaffer, N. (Eds.). (2008). *Game Usability: Advice From The Experts For Advancing The Player Experience*. San Francisco, CA: Morgan Kauffman.

Jacobsen, N. E., Hertzum, M., & John, B. E. (1998). The evaluator effect in usability tests. In *CHI 98 conference summary on Human factors in computing systems* (pp. 255-256). Los Angeles, CA: ACM.

Järvinen, A., Heliö, S., & Mäyrä, F. (2002). *Communication and Community in Digital Entertainment Services*. University of Tampere.

Johnson, D., & Wiles, J. (2003). Effective affective user interface design in games. *Ergonomics*, *46*(13), 1332–1345. doi:10.1080/00140130310001610865

Köffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A., & Tscheligi, M. (2010). Using Heuristics to Evaluate the Overall User Experience of Video Games and Advanced Interaction Games. In Bernhaupt, R. (Ed.), *Evaluating User Experience in Games: Concepts and Methods*. Springer. doi:10.1007/978-1-84882-963-3_13

Koivisto, E. M. I., & Korhonen, H. (2006). *Mobile Game Playability Heuristics*. Retrieved September 10, 2010, from http://www.forum.nokia.com.

Korhonen, H., & Koivisto, E. M. I. (2006). Playability heuristics for mobile games. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 9-16). Helsinki, Finland: ACM. doi:10.1145/1152215.1152218

Korhonen, H., & Koivisto, E. M. I. (2007). Playability heuristics for mobile multi-player games. In *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts* (pp. 28-35). Perth, Australia: ACM.

Laitinen, S. (2006). Do Usability Expert Evaluation and Testing Provide Novel and Useful Data For Game Development? *Journal of Usability Studies*, *2*(1), 64–75.

Ling, C., & Salvendy, G. (2005). Extension of Heuristic Evaluation Method: a Review and Reappraisal. Ergonomia. *An International Journal of Ergonomics and Human Factors*, *27*(3), 179–197.

Malone, T. W. (1982). Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 conference on Human factors in computing systems* (pp. 63-68). Gaithersburg, MD: ACM.

Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S., & Ames, M. (2003). Heuristic evaluation of ambient displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 169-176). Ft. Lauderdale, FL: ACM.

Mäyrä, F., & Ermi, L. (2005). Fundamental Components of the Gameplay Experience: Analysing Immersion. Retrieved September 15, 2010, from http://www.digra.org/dl/db/ 06276.41516.pdf

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. doi:10.1037/h0043158

Molich, R., & Dumas, J. S. (2008). Comparative Usability Evaluation (CUE-4). *Behaviour & Information Technology*, *27*, 263–281. doi:10.1080/01449290600959062

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 373-380). Monterey, CA: ACM.

Nielsen, J. (1994a). Heuristic Evaluation. In Nielsen, J., & Mack, R. L. (Eds.), *Usability Inspection Methods*. New York: John Wiley & Sons.

Nielsen, J. (1994b). *Usability Engineering*. San Francisco, CA: Morgan Kaufmann.

Nielsen, J. (2005a). Heuristic Evaluation. *UseIt website*. Retrieved September 10, 2010, from http://www.useit.com/ papers/heuristic/

Nielsen, J. (2005b). How to Conduct a Heuristic Evaluation. *UseIt website*. Retrieved September 10, 2010, from http://www.useit.com/papers/heuristic/heuristic_evaluation.html

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people* (pp. 249-256). Seattle, WA: ACM.

Pagulayan, R., & Steury, K. (2004). Beyond usability in games. *interactions, 11*(5), 70-71.

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2008). User-centered Design in Games. In Jacko. J Sears A., (Eds.). *Handbook for Human-Computer Interaction in Interactive Systems*, Second Edition, Lawrence Erlbaum Associates, Inc., 741-759.

Pinelle, D., Wong, N., & Stach, T. (2008a). Heuristic evaluation for games: usability principles for video game design. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (pp. 1453-1462). Florence, Italy: ACM.

Pinelle, D., Wong, N., & Stach, T. (2008b). Using genres to customize usability evaluations of video games. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (pp. 129-136). Toronto, Ontario, Canada: ACM.

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009). Usability heuristics for networked multiplayer games. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 169-178). Sanibel Island, FL: ACM.

PurestProdigy. (2007). The Simpsons: Minutes to Meltdown. *GameFaqs user review*. Retrieved September 10, 2010, from http://www.gamefaqs.com/mobile /mobile/review/R125801.html

Schaffer, N. (2007). *Heuristics for Usability in Games*. White Paper. Retrieved September 10, 2010, from http://www.playerfriendly.com/ files/heuristics.pdf

Sim, G., Read, J., & Cockton, G. (2009). Evidence Based Design of Heuristics for Computer Assisted Assessment. In T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. Prates, & M. Winckler (Eds.), *Human-Computer Interaction – INTERACT 2009*, Lecture Notes in Computer Science (Vol. 5726, pp. 204-216). Springer Berlin / Heidelberg. Retrieved September 15, 2010, from http://dx.doi.org/10.1007/978-3-642-03655-2_25

Snow, B. (2007). Game Usability 101. *BusinessWeek*. Retrieved September 10, 2010, from http://www.businessweek.com/ innovate/content/oct2007 /id20071012_041625.htm.

Zuk, T., Schlesier, L., Neumann, P., Hancock, M. S., & Carpendale, S. (2006). Heuristics for information visualization evaluation. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization* (pp. 1-6). Venice, Italy: ACM.

ADDITIONAL READING

Adams, E. (2009). *Fundamentals of Game Design* (2nd ed.). Berkeley, CA: New Riders Press.

Barr, P., Noble, J., & Biddle, R. (2007). Video game values: Human-computer interaction and games. *Interacting with Computers*, *19*(2), 180–195. doi:10.1016/j.intcom.2006.08.008

Bekker, M. M., Baauw, E., & Barendregt, W. (2008). A comparison of two analytical evaluation methods for educational computer games for young children. *Cognition Technology and Work*, *10*(2), 129–140. doi:10.1007/s10111-007-0068-x

Benford, S., Crabtree, A., Reeves, S., Sheridan, J., Dix, A., Flintham, M., & Drozd, A. (2006). The Frame of the Game: Blurring the Boundary between Fiction and Reality in Mobile Experiences. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 427-436). Montréal, Québec, Canada: ACM.

Bertini, E., Gabrielli, S., & Kimani, S. (2006). Appropriating and assessing heuristics for mobile computing. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 119-126). Venezia, Italy: ACM.

Bond, M., & Beale, R. (2009). What makes a good game?: using reviews to inform design. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (pp. 418-422). Cambridge, United Kingdom: British Computer Society. Retrieved September 15, 2010, from http://portal.acm.org/ citation.cfm?id=1671065&dl=GUIDE&coll=GUIDE&CFID=104687055&CFTOKEN=98301315

Bosser, A., & Nakatsu, R. (2006). Hardcore Gamers and Casual Gamers Playing Online Together. In R. Harper, M. Rauterberg, & M. Combetto (Eds.), *Entertainment Computing - ICEC 2006*, Lecture Notes in Computer Science (Vol. 4161, pp. 374-377). Springer Berlin / Heidelberg. Retrieved September 15, 2010, from http://dx.doi.org/10.1007/ 11872320_53

Crawford, C. (2003). *Chris Crawford on Game Design*. Berkeley, CA: New Riders Games.

De Kort, Y. A. W., & Ijsselsteijn, W. A. (2008). People, places, and play: player experience in a socio-spatial context. *Computers in Entertainment*, *6*(2), 1–11. doi:10.1145/1371216.1371221

Dongseong, C., Kim, H., & Kim, J. (1999). Toward the Construction of Fun Computer Games: Relative Importance of Design Factors between Developers versus Players. *Personal and Ubiquitous Computing*, *3*(3), 92–104.

Dumas, J. S., & Loring, B. A. (2008). *Moderating Usability Tests: Principles and Practices for Interacting*. San Francisco, CA: Morgan Kaufmann.

Dumas, J. S., Molich, R., & Jeffries, R. (2004). Describing usability problems: are we sending the right message? *Interaction*, *11*(4), 24–29. doi:10.1145/1005261.1005274

Falstein, N. (2005). Understanding Fun - The Theory of Natural Funativity. In S. Rabin (Ed.), *Introduction to Game Development* (Game Developement Series, pp. 71-97). Hingham, MA: Charles River Media, Inc.

Febretti, A., & Garzotto, F. (2009). Usability, playability, and long-term engagement in computer games. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4063-4068). Boston, MA: ACM.

Hornbæk, K., & Frøkjær, E. (2008). Comparison of techniques for matching of usability problem descriptions. *Interacting with Computers*, *20*(6), 505–514. doi:10.1016/j.intcom.2008.08.005

Hvannberg, E. T., Law, E. L., & Lárusdóttir, M. K. (2007). Heuristic evaluation: Comparing ways of finding and reporting usability problems. *Interacting with Computers*, *19*(2), 225–240. doi:10.1016/j.intcom.2006.10.001

Jegers, K. (2008). Investigating the Applicability of Usability and Playability Heuristics for Evaluation of Pervasive Games. In *2008 Third International Conference on Internet and Web Applications and Services* (pp. 656-661). Presented at the 2008 3rd International Conference on internet and Web Applications and Services (ICIW), Athens, Greece.

Kampmann, W. (2003). Playing and Gaming - Reflections and Classifications. *Game Studies*, 3(1). Retrieved September 15, 2010, from http://www.gamestudies.org/ 0301/walther/

Korhonen, H. (2010). *Comparison of Playtesting and Expert Review Methods in Mobile Game Evaluation*. Paper presented at the International Conference on Fun and Games, 2010.

Korhonen, H., Montola, M., & Arrasvuori, J. (2009). *Understanding Playful User Experience through Digital Games*. Paper presented at the Designing Pleasurable Products and Interfaces.

Ling, C., & Salvendy, G. (2009). Effect of evaluators' cognitive style on heuristic evaluation: Field dependent and field independent evaluators. *International Journal of Human-Computer Studies*, *67*(4), 382–393. doi:10.1016/j.ijhcs.2008.11.002

Nielsen, J. (1992). Reliability of severity estimates for usability problems found by heuristic evaluation. In *Posters and short talks of the 1992 SIGCHI conference on Human factors in computing systems* (pp. 129-130). Monterey, California: ACM.

Omar, H. M., & Jaafar, A. (2008). Playability Heuristics Evaluation (PHE) approach for Malaysian educational games. In *2008 International Symposium on Information Technology* (pp. 1-7). Presented at the 2008 International Symposium on Information Technology, Kuala Lumpur, Malaysia.

Orvis, K. A., Horn, D. B., & Belanich, J. (2008). The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. *Computers in Human Behavior*, *24*(5), 2415–2433. doi:10.1016/j.chb.2008.02.016

Paavilainen, J. (2010). *Critical Review on Video Game Evaluation Heuristics: Social Games Perspective*. Paper presented at the international conference on Future Play: Research, Play, Share.

Papaloukas, S., Patriarcheas, K., & Xenos, M. (2009). Usability Assessment Heuristics in New Genre Videogames. In *Proceedings of the 2009 13th Panhellenic Conference on Informatics* (pp. 202-206). IEEE Computer Society. Retrieved September 15, 2010, from http://portal.acm.org/citation.cfm?id=1684685

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009). Usability heuristics for networked multiplayer games. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 169-178). Sanibel Island, FL: ACM.

Qin, H., Rau, P. P., & Salvendy, G. (2010). Effects of different scenarios of game difficulty on player immersion. *Interacting with Computers*, *22*(3), 230–239. doi:10.1016/j.intcom.2009.12.004

Rollings, A., & Adams, E. (2003). *Andrew Rollings and Ernest Adams on Game Design (Ltd Rmst.)*. Berkeley, CA: New Riders Games.

Rouse, R. (2001). *Game Design: Theory and Practice*. Plano, TX: Wordware Publishing.

Rubin, J. (1994). *Handbook of Usability Testing: How to plan, design and conduct effective test*. John Wiley & Sons.

Salen, K., & Zimmerman, E. (2003). *Rules of Play: Game Design Fundamentals*. Cambridge, MA: The MIT Press.

Schell, J. (2008). *The Art of Game Design*. San Francisco, CA: Morgan Kaufmann.

Sutcliffe, A., & Gault, B. (2004). Heuristic evaluation of virtual reality applications. *Interacting with Computers*, *16*(4), 831–849. doi:10.1016/j.intcom.2004.05.001

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment, 3*(3).

Wiberg, C. (2005). *Fun in the Home: Guidelines for Evaluating Interactive Entertainment on the Web*. Paper presented at the Conference on Human-Computer Interaction International.

Wiberg, C., Jegers, K., & Desurvire, H. (2009). How Applicable is Your Evaluation Methods - Really? Analysis and Re-design of Evaluation Methods for Fun and Entertainment. In *Proceedings of the 2009 Second International Conferences on Advances in Computer-Human Interactions* (pp. 324-328). IEEE Computer Society. Retrieved September 15, 2010, from http://portal.acm.org/ citation.cfm?id=1509869.1509931&coll=Portal&dl=GUIDE&CFID=104691490&CFTOKEN =32024841

KEY TERMS AND DEFINITIONS

Expert Review: An analytical evaluation method in which experts conduct an evaluation. The experts have good knowledge of usability/playability principles and they are preferably also experts in the domain.

Game Evaluation: Game evaluations assess playability or game experience of a game by using different evaluation methods. The methods can be either analytical evaluation methods or user testing.

Game Experience: An experience enabled by the game for the player. Usability and playability are game centric terms where as game experience is related to the experiential engagement between the player and the game.

Heuristics: Heuristics are guidelines, rule of thumb statements, which reflect the desirable aspects of a given product.

Playability: Playability is defined as aspects that relate to desirable aspects of a good game. The game has good playability when the user interface is intuitive and unobtrusive, so that the player can concentrate on playing the game. Gameplay includes, for example, game mechanics, narrative, and goals that the player tries to achieve. Fun and challenge are created by the gameplay; the game has good playability especially when the gameplay is understandable, balanced, suitably difficult, and engaging. In multiplayer games, players' social interaction also affects playability.

Usability: The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. Additionally, other factors such as learnability, memorability and error prevention can be also considered to be part of usability.

Video Game: A type of game existing as and controlled by software, run by a device with video terminal and played with an interaction interface.

ENDNOTES

1 http://www.usabilityfirst.com/glossary/term_657.txl
2 http://mgsuserresearch.com/
3 http://www.eamobile.com/Web/mobile-games/the-simpsons-minutes-to-meltdown

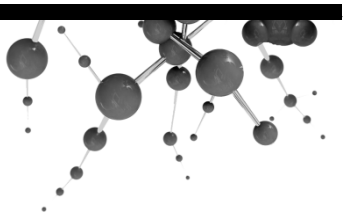
APPENDIX 1. EVALUATION HEURISTICS BY DESURVIRE ET AL. (2004)

ID	Category	Description
GP 1	Game Play	Player's fatigue is minimized by varying activities and pacing during game play.
GP 2	Game Play	Provide consistency between the game elements and the overarching setting and story to suspend disbelief.
GP 3	Game Play	Provide clear goals, present overriding goal early as well as short-term goals throughout play.
GP 4	Game Play	There is an interesting and absorbing tutorial that mimics the game play.
GP 5	Game Play	The game is enjoyable to replay.
GP 6	Game Play	Game should be balanced with multiple ways to win.
GP 7	Game Play	Player is taught skills early that you expect the players to use later, or right before the new skill is needed.
GP 8	Game Play	Players discover the story as part of game play.
GP 9	Game Play	Even if the game cannot be modeless, it should be perceived as modeless.
GP 10	Game Play	The game is fun for the player first, the designer second and the computer third. That is, if the non-expert player's experience isn't put first, excellent game mechanics and graphics programming triumphs are meaningless.
GP 11	Game Play	Player should not experience being penalized repetitively for the same failure.
GP 12	Game Play	Player's should perceive a sense of control and impact onto the game world. The game world reacts to the player and remembers their passage through it. Changes the player makes in the game world are persistent and noticeable if they back-track to where they've been before.
GP 13	Game Play	The first player action is painfully obvious and should result in immediate positive feedback.
GP 14	Game Play	The game should give rewards that immerse the player more deeply in the game by increasing their capabilities (power-up), and expanding their ability to customize.
GP 15	Game Play	Pace the game to apply pressure but not frustrate the player. Vary the difficulty level so that the player has greater challenges as they develop mastery. Easy to learn, hard to master.
GP 16	Game Play	Challenges are positive game experiences, rather than negative experience (results in their wanting to play more, rather than quitting).
GS 1	Game Story	Player understands the storyline as a single consistent vision.
GS 2	Game Story	Player is interested in the storyline. The story experience relates to their real life and grabs their interest.
GS 3	Game Story	The player spends time thinking about possible story outcomes.
GS 4	Game Story	The player feels as though the world is going on whether their character is there or not.
GS 5	Game Story	The player has a sense of control over their character and is able to use tactics and strategies.
GS 6	Game Story	Player experiences fairness of outcomes.
GS 7	Game Story	The game transports the player into a level of personal involvement emotionally (e.g., scare, threat, thrill, reward, punishment) and viscerally (e.g., sounds of environment).
GS 8	Game Story	Player is interested in the character because (1) they are like me, (2) they are interesting to me, (3) the characters develop as action occurs.
GM 1	Game Mechanics	Game should react in a consistent, challenging, and exciting way to the player's actions (e.g., appropriate music with the action).
GM 2	Game Mechanics	Make effects of the Artificial Intelligence (AI) clearly visible to the player by ensuring they are consistent with the player's reasonable expectations of the AI actor.
GM 3	Game Mechanics	A player should always be able to identify their score/status in the game.
GM 4	Game Mechanics	Mechanics/controller actions have consistently mapped and learnable responses.
GM 5	Game Mechanics	Shorten the learning curve by following the trends set by the gaming industry to meet user's expectations.
GM 6	Game Mechanics	Controls should be intuitive and mapped in a natural way; they should be customizable and default to industry standard settings.
GM 7	Game Mechanics	Player should be given controls that are basic enough to learn quickly yet expandable for advanced options.

ID	Category	Description
GU 1	Game Usability	Provide immediate feedback for user actions.
GU 2	Game Usability	The player can easily turn the game off and on, and be able to save games in different states.
GU 3	Game Usability	The player experiences the user interface as consistent (in control, color, typography, and dialog design) but the game play is varied.
GU 4	Game Usability	The player should experience the menu as part of the game.
GU 5	Game Usability	Upon initially turning the game on the player has enough information to start playing.
GU 6	Game Usability	Players should be given context-sensitive help while playing so that they do not get stuck or have to rely on a manual.
GU 7	Game Usability	Sounds from the game provide meaningful feedback or stir a particular emotion.
GU 8	Game Usability	Players do not need to use a manual to play the game.
GU 9	Game Usability	The interface should be as non-intrusive to the player as possible.
GU 10	Game Usability	Make the menu layers well-organized and minimalist to the extent that the menu options are intuitive.
GU 11	Game Usability	Get the player involved quickly and easily with tutorials and/or progressive or adjustable difficulty levels.
GU 12	Game Usability	Art should be recognizable to the player, and speak to its function.

APPENDIX 2. EVALUATION HEURISTICS BY KORHONEN AND KOIVISTO (2006)

ID	Category	Description
GU 1	Game Usability	Audiovisual representation supports the game.
GU 2	Game Usability	Screen layout is efficient and visually pleasing.
GU 3	Game Usability	Device UI and game UI are used for their own purpose.
GU 4	Game Usability	Indicators are visible.
GU 5	Game Usability	The player understands terminology.
GU 6	Game Usability	Navigation is consistent, logical and minimalist.
GU 7	Game Usability	Control keys are consistent and follow standard conventions.
GU 8	Game Usability	Game controls are convenient and flexible.
GU 9	Game Usability	The game gives feedback on the player's actions.
GU 10	Game Usability	The player cannot make irreversible errors.
GU 11	Game Usability	The player does not have to memorize things unnecessarily.
GU 12	Game Usability	The game contains help.
MO 1	Game Mobility	The game and play sessions can be started quickly.
MO 2	Game Mobility	The game accommodates to the surroundings.
MO 3	Game Mobility	Interruptions are handled reasonably.
GP 1	Gameplay	The game provides clear goals or supports player created goals.
GP 2	Gameplay	The player sees the progress in the game and can compare the results.
GP 3	Gameplay	The players are rewarded and rewards are meaningful.
GP 4	Gameplay	The player is in control.
GP 5	Gameplay	Challenge, strategy, and pace are in balance.
GP 6	Gameplay	The first-time experience is encouraging.
GP 7	Gameplay	The game-story supports the gameplay and is meaningful.
GP 8	Gameplay	There are no repetitive or boring tasks.
GP 9	Gameplay	The players can express themselves.
GP 10	Gameplay	The game supports different playing styles.
GP 11	Gameplay	The game does not stagnate.
GP 12	Gameplay	The game is consistent.
GP 13	Gameplay	The game uses orthogonal unit differentiation (units in the game should be designed so that they are functionally different)
GP 14	Gameplay	The player does not lose any hard won possessions.



Paper V

Korhonen, H. (2010). Comparison of Playtesting and Expert Review Methods in Mobile Game Evaluation. In *Proceedings of the International Conference on Fun and Games (Fun'n'Games 2010, Leuven, Belgium),* 18-27. New York, NY, USA: ACM. doi: 10.1145/1823818.1823820

Comparison of Playtesting and Expert Review Methods in Mobile Game Evaluation

Hannu Korhonen
Nokia Research
P.O. Box 1000
33721 Tampere, Finland
hannu.j.korhonen@nokia.com

ABSTRACT
Selecting an evaluation method for product evaluations depends on many issues, such as the development stage of the product, time schedule, resources, and money that can be invested on the evaluation. The user testing and expert review methods are probably the most common ones when productivity software is being evaluated. Conducting a playtesting with players is commonly used by game designers, but the expert review method has not received that much attention, although it has proven to be an efficient and useful method. In this paper, we present a comparison study of the playtesting and expert review methods in mobile game evaluation. Our objective is to compare the effectiveness of the expert review method with playtesting. Results indicate that the expert review method is able to identify playability problems as accurately as playtesting, but in addition, it identifies problems that are crucial for the playability of the game.

Categories and Subject Descriptors
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces, *Evaluation/methodology*.

General Terms
Measurement, Experimentation, Human Factors

Keywords
Evaluation method, Expert Review, Playtesting, Playability Heuristics, Comparison Study.

1. INTRODUCTION
HCI researchers have developed multiple evaluation methods for testing the usability of productivity software. Selecting a method is not only about the method itself, but it also depends on other attributes such as resources, time, and money that can be invested on the evaluation. There are basically two mainstreams of evaluation methods: the user testing and inspection methods. In user testing, persons from a target user group interact with a product, and their behavior and experiences are collected either automatically or by observers. Even though this method enables developers to see how the product is used and perceived by the target users, arranging a user testing is time and resource consuming [25].

There are multiple inspection methods available that can be used for product evaluations. Nielsen calls these inspection methods discount methods, since they do not usually include users from the target user group, but the evaluation is conducted by experts [13]. The most popular inspection method is the expert review method (a.k.a. Heuristic Evaluation) developed by Nielsen and Molich [20]. In this method, usability experts evaluate a product by using usability heuristics (e.g. [18]) as principles to check the user interface of the product. The benefit of the expert review method is that it is fast, and knowledgeable experts can conduct an evaluation within couple of hours by using functional prototypes, low-fidelity prototypes (i.e. paper mockups) or even concept and interaction descriptions.

Despite its popularity in productivity software evaluations, the expert review method has not received much attention among game designers, or at least the usage of the method has not been reported extensively. In game design literature, focus groups and playtesting are mentioned frequently as evaluation methods for evaluating game designs [7], [24], [26]. Recently, game researchers have started to develop the expert review method for game evaluations [3], [6], [10], [23], and their focus has been in developing required heuristics. Comparing different evaluation methods in game evaluations has not been studied extensively.

Games differ in many respects from productivity software, which should be taken into account when applying evaluation methods developed for productivity software evaluations to game evaluations. Pagulayan et al. describe differences between games and productivity software [14]. According to them, productivity applications are tools, and the design intention is to make tasks easier, efficient, less error-prone, and increase the quality of the results. Games, instead, are intended to be pleasurable to play and sufficiently challenging in order to provide a good gaming experience. In addition, learning the goals, strategies and tactics to succeed in a game is part of the fun [14]. The intention of the game evaluation is to reduce the obstacles of fun, rather than the obstacles of accomplishments. This is a remarkable difference compared to productivity software evaluation, which focuses on the efficiency and ease of use of the product.

The contribution of the paper to the game research domain is that it provides empirical results on how the expert review method results correspond to playtesting results. In this study, we evaluated a mobile game by using two methods, and the results indicate that the expert review method very accurately identifies

Fun and Games 2010, September 15-17, 2010, Leuven, Belgium.

the playability problems that came up in playtesting. Therefore, using the expert review method should be included in the evaluation method toolbox of game designs, and could be used to evaluate playability problems in games.

2. RELATED WORK

In this section, we review some studies that have applied the user testing and expert review methods to evaluating productivity software or web pages, and have compared evaluation results. Then we take a look at published game evaluations that have used several evaluation methods to evaluate digital games.

2.1 Comparative Productivity Software Evaluations

Karat [17], (p.204) describes that the number of identified usability problems and their type is one of the issues that define the usefulness of an evaluation method when comparing different evaluation methods. Molich et al. have conducted a series of comparative usability evaluations to find out what kinds of usability problems are found in the same application or web page [14]. In these studies, voluntary usability teams plan and conduct an evaluation of an application or web site by using the method they prefer and report results. Molich et al. found that the correspondence of the results is not very high, since 75% of the identified usability problems were reported only by a single team [14]. Doubleday et al. have reported that in their study, 39 % of usability problems that were identified in the user testing were not discovered in the heuristic evaluation [2].

Desurvire [17], (p.175) states that if the majority of identified problems are minor problems and there are more problems identified than in the laboratory test, an inspection method should not be considered very useful. The results from Jeffries et al. [5] showed that the heuristic evaluation method identified the most usability problems and more of the serious problems. However, the results of the study are inconclusive because different evaluation methods seemed to identify different usability problems. The overlap of the identified problems was only 10-15 percent between the methods [17], (p.210). Other researchers have reported similar results when evaluating productivity software. Desurvire reported that in their study, usability experts found 44 percent of usability problems that also occurred in the usability testing [17], (p.185). Karat reported that in their study, a third of the significant usability problems were identified with all methods. Gray and Salzman have criticized that many of these older studies contain errors in the experiment design, resulting in the fact that conclusions drawn from the data are not reliable or valid [8]. They state that the experiment should be designed very carefully to achieve reliable and comparable results.

There are at least three reasons to why evaluation results are different. First, in some studies, the evaluated product was not the same. Desurvire et al. reported that the user testing was conducted with the real system, whereas usability experts evaluated a set of flowchart diagrams of the system [17], (p.181). Another reason is that tasks or scenarios that are used in the evaluation are not the same. Molich et al. allowed each evaluation team to develop their own set of tasks, and the result was that almost half of the tasks were unique and used only by one team [14]. The third reason for different results is the "evaluator effect" [4]. Nielsen and Molich also reported that, regardless of the same background, the correlation between the numbers of usability problems identified by individual evaluators was not high. Nielsen [13] (p.59) concluded that it is preferable to use usability specialists to conduct evaluations, and for optimal performance, they should be double experts[1]. Recruiting double experts to conduct an evaluation is not an easy task, and therefore, it is a tempting idea to use actual end users as evaluators. However, research results indicate that this does not work, since the users do not have the required knowledge and understanding of usability principles [13] (p.59). Muller et al. propose the Participatory Heuristic Evaluation method, in which end users participate in the heuristic evaluation [10]. This is reasonable when end users are members of the development team.

Recent studies have indicated that both user testing and expert review can provide similar kinds of results. Molich and Dumas have studied how the usability testing and expert review methods compare. The results from this study show that the expert review method can provide valuable results in terms of identified problems when compared to user testing [13].

2.2 Game Evaluations

User testing or playtesting is the most popular evaluation method, and it is also described as the main evaluation method for game designers in game design literature [7], [24], [26]. Pagulayan et al. have conducted playtesting by using open-ended tasks. The purpose of these tests is to gather data on how players prioritize tasks and goals in the game [14]. They can also reveal how players understand the game mechanics or controls in the game.

Analytical evaluation methods and their potential applications to game evaluations have also been studied. The main focus of these studies has been to develop heuristics that are used during the evaluation [3], [10], [23]. More recently, a comparison study for evaluating two heuristic sets has also been conducted [11]. Baauw et al. have introduced an analytical expert evaluation method called the Structured Expert Evaluation Method (SEEM). Instead of using heuristics, the method includes a set of questions based on previous work from Norman and Malone [2]. This method has been used to evaluate games for children.

There is also increasing interest towards using automatic recording of interaction with the game systems to evaluate playability issues and gaming experience. Kim et al. have introduced the TRUE system which records data streams combining events of interest, contextual information and subjective opinions of the players to get a holistic view of what is happening in the game and possible problems the players have in playing the game [9]. Drachen and Canossa have also presented an instrumentation system that is used at IO Interactive [4]. This system collects various data from the play session. These kinds of systems are excellent tools for gathering very detailed information about player behavior in the game.

Studies on the comparison of different methods in game evaluations are very minimal. Laitinen has conducted a usability expert evaluation and user testing in a case study of a computer game [8]. The expert evaluation used traditional usability heuristics originally defined by Nielsen and Molich [12]. In the user testing, moderators introduced the game to the players and

[1] The evaluator has knowledge of general usability principles as well as product domain and task flows of the end users.

explained the background story and the starting point of the game. Players played the game for one and the half hours, and for the last 15 minutes, a cheat mode was activated in order to evaluate features that were not directly available for the players. The results do not directly indicate how well the results of the evaluation methods correspond to each other. However, Laitinen concluded that both methods provide useful data and identified problems in the design that were new to game developers [12].

Desurvire et al. performed a comparative evaluation of a computer game using both heuristic evaluation and playtesting [3]. Heuristics that they used in the inspection were specifically designed for game evaluations. One playability expert performed the heuristic evaluation. The two-hour playtesting was arranged with four players. The evaluation was conducted with an early prototype of the game. Results of the study indicate that both methods identify similar kinds of playability problems. The playtesting identified specific problems in the interface, whereas the expert evaluation identified general interface design issues.

3. The Study

The objective of the study was to compare the playtesting and expert review methods and to see how much evaluation results differ when they are used to evaluate a mobile game. The expert review was conducted before the playtesting. The author analyzed evaluation data after both evaluations were finished.

3.1 The Mobile Game

The mobile game was a commercial 3D action/puzzle game that can be played on smart phones. The game was a new version of one of the first games that started gaming on mobile phones. Therefore, evaluators and players who participated in the playtesting were familiar with the game concept. The evaluators and four out of six players have played the previous version of the game. The game was still under development, but the evaluated version was launched for beta testing, which means that all features of the game were implemented and it could be played on a target device [15]. This allowed the game to be evaluated thoroughly and the evaluation sessions were realistic. Both evaluations were conducted by using the same version of the game. The evaluators and the players used Nokia N73 mobile phones to play the game.

3.2 The Expert Review of the Game

The expert review was conducted by two playability experts and it followed recommended procedure [21], [19]. The evaluators were selected based on their expertise in conducting game evaluations and productivity software evaluations. In addition, both the evaluators play different mobile games regularly, which gave them expertise with mobile games. They were also familiar with the playability heuristics (See Appendix A) that are specifically designed for evaluating mobile games [10].

The evaluators did not belong to the development team of the game, and they did not have any previous experience with the game prior to the evaluation session. This provided a realistic context for the evaluators, and it resembled the situation that players will face when they get a new game for their device. The evaluators were instructed both to explore the user interface of the game and to try to complete as many levels as they could during the evaluation session. The instructions did not include tasks or scenarios that the evaluators should follow, but they were free to explore the game as they liked.

The evaluation started from the moment when the evaluators launched the game for the first time. First, the evaluators examined the game menu and general settings of the game and walked through the first levels that served as a tutorial for the game. Identified playability problems were written down briefly and the violated heuristic was assigned. The purpose was to keep the paper work of the evaluation to a minimum during the first moments, because it would otherwise disturb the evaluators' gaming experience. However, it is extremely important to record these first impressions with the game, because a player will learn and adapt quickly to design problems and valuable information is lost, if it is not recorded immediately.

As the evaluation continued, the evaluators focused thoroughly on both game usability and gameplay issues of the game. Sometimes the evaluators needed to play some levels several times before they could complete them. However, this also allowed them to explore the possibilities of the game and try out different strategies and playing styles.

The evaluation continued until the time required for finding new playability issues increased dramatically. This was based on the evaluators' own judgment. After that, the evaluators walked through identified playability problems one by one and discussed about their findings. The identified playability problems were clarified and duplicates were removed from a combined list. Finally, a violated heuristic and severity of the playability problem were assigned. Recommendations to fix the playability problems were documented as a final step of the evaluation.

The expert review session took approximately 3-4 hours including the evaluation of the game, discussion between the evaluators, and documenting the findings. Time reserved for playing the game was approximately one hour. The evaluators did not finish the whole game, but they completed the first eight levels.

3.3 The Playtesting of the Game

The playtesting was conducted in a usability laboratory and one participant was playing the game at a time. Six players were recruited for the evaluation. One participant out of the six was female. The average age of the players was 30 years, ranging from 26 to 35. All participants were experienced mobile phone users and they had played mobile games to some extent. Only two players played mobile games frequently. Others can be categorized as casual players of mobile games, which was the target population of the evaluated game.

The procedure of the playtesting followed the standard procedure of user testing (e.g. [5], [25]). In the beginning, the moderator instructed a player on how to think aloud during the session and collected background information. After that, the player was allowed to start playing the game. The moderator observed the game session and how the player played the game. The moderator also asked questions to verify his observations during the evaluation session. There was also another observer in an observation room to make notes from the game session. The session lasted 60-90 minutes including the introduction phase, playing the game, and a post-test interview. The time for playing the game was approximately 60 minutes, and the difference in the total time was caused by the length of the post-test interview which consisted of open-ended questions and a questionnaire.

Although the playtesting was similar to standard user testing, there was one significant difference. Instead of using specific or predefined tasks, the players were instructed with a single open-ended task which was formulated as follows:

"Play the game as you would play it on your own. The moderator will ask you questions and tell you when to stop."

The same method is also used in game evaluations at Microsoft Games User-Testing Group [22]. The open-ended task does not instruct a player to perform or achieve anything in particular, or to play the game in a certain way. Instead, it allows the moderator to observe whether a player understands the goals and other aspects in the game. In our opinion, this makes the evaluation session more realistic and it also corresponds to the situation where the experts were playing the game.

The playtesting sessions were video recorded for later analysis. The recording system consisted of two cameras and a video mixer. One camera was recording the facial expressions and body movements of the participant, and another camera was mounted to a mobile phone to capture events of the screen and interaction with a keypad (Figure 1). During the sessions, the video stream was also transmitted to a monitor located in front of the moderator. This enabled the moderator to have a clear visibility of the content on the screen and observe the player's actions without disturbing them. Having the moderator next to the player also allowed for natural conversation during the test session.

Figure 1. A mini camera mounted to a mobile phone (The mobile phone differs from the actual device used during the playtesting)

3.4 The combined analysis of the evaluation results

Our main research question for the study was how well an expert evaluation and playtesting can identify the same playability problems in the game. In order to do this, we harmonized findings from both evaluations. In the analysis phase, all playability problems were analyzed one by one and duplicates were removed. We then categorized the remaining problems according to our playability heuristics and severity rankings. Finally, results from both evaluations were cross-checked to ensure that the same playability problems were recorded equally.

4. Results

In this section, we describe results from both the expert review and playtesting and go through the differences that were found from the results.

4.1 Playability problems

The number of playability problems reported by the two evaluation methods was quite similar. Six playtesting sessions reported 46 playability problems altogether, whereas the combined list of playability problems from two experts contained 32 playability problems. There were both common problems and unique problems reported by a single method.

The combined list of playability problems from both evaluations contains 53 playability problems. We did not include playability problems violating mobility heuristics in the analysis, because those were reported only by the two playability experts. Evaluating mobility aspects during the playtesting would have made the test sessions more complicated. In addition, we excluded all positive observations from the analysis, as those were only reported by the experts and they do not contribute to the problem identification of the game. Hence, the analysis contains playability issues related to user interface and gameplay of the game.

Both evaluations were heavily user interface oriented, as 70% of the reported playability problems concerned user interface issues and 30% of the problems were identified as gameplay problems. Figure 2 shows problems categorized by the playability heuristics they violated. An interesting observation from the reported problems is that both evaluation methods reported 20 user interface problems, and in addition, playtesting reported 17 user interface problems. The expert review did not report any user interface related playability problems that would not have been reported by the playtesting.

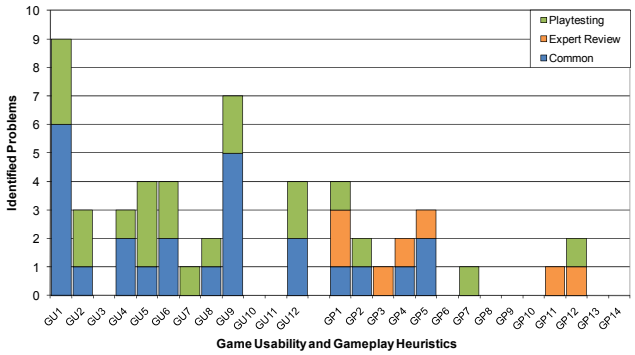


Figure 2. Playability problems categorized by violated heuristics

When looking at playability problems related to gameplay, we can see some differences in what kinds of problems were reported. There are 16 playability problems in total. Five playability problems are reported by both methods, seven problems are reported by two playability experts, and four problems are reported in the playtesting (Table 1).

Table 1. The number of playability problems and their severity identified by different evaluation methods

	Game Usability			Gameplay			
	Major	Medium	Minor	Major	Medium	Minor	Total
Common	9	10	1	3	2	0	25
Experts Only	0	0	0	3	4	0	7
Players Only	0	7	10	1	2	1	21
							53

Although 47% of the playability problems of the game were reported by both evaluation methods, there seem to be lots of playability problems that were reported independently from the other method (Figure 3). This provides an interesting starting point for viewing the differences between the results in more detail, as well as possible causes of these differences.

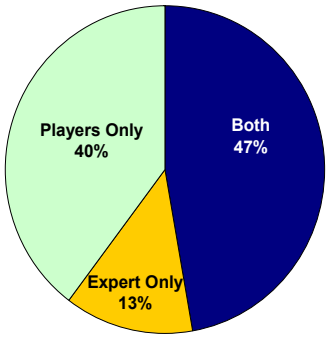


Figure 3. Playability problems identified by different methods

4.2 Playtesting Findings

Both evaluation methods focused on user interface issues, and 70% of the reported problems belonged to this category. The playtesting reported several user interface problems that were not discovered by the playability experts. When looking at these problems in detail, the first observation is that the majority of these problems seem to be quite player specific. The playtesting reported 21 playability problems in total that only players encountered (Table 1). However, in 81% of the cases (17 problems) only one out of six players encountered the problem. Furthermore, the problems did not seem to be severe, since the severity ranking of the problems was either medium or minor. 10 problems were rated as minor problems, indicating that someone finds them disturbing, but they do not really affect the player experience. Seven problems were rated medium, meaning that players initially find them disturbing, but once the players learned them, they are not problematic anymore. Examples of such playability problems are visualization and terminology issues and difficulties with control keys.

However, there were a couple of playability problems that were reported by several players. Since there were only six players participating to the playtesting, a problem reported by more than one player can be a significant issue and problematic to a larger player population as well. Two of them were ranked minor problems, whereas one problem had medium severity.

Two players had problems with in-game instructions, because they did not explain the purpose of some special items in the game clearly enough and the items were introduced to the players too early. The players tried to find these items on basic levels, even though they were available only on advanced levels. The playability experts did not consider this as a problem.

Three players had difficulties in understanding one visualization effect that was used in the game. When the player completed the level or a game character died, the screen went into grayscale mode. The players misinterpreted this to be some sort of special mode, because the game character was still moving in the game world and the player could control the character awhile. The playability experts did not point out this specific problem, because there were other problems affecting in this situation at the same time. The problem was ranked with minor severity.

The most severe playability problem concerning the user interface was related to indicators that were used to present essential game information to the player. Five out of six players had difficulties in recognizing indicators for the time limit, level progress and special abilities indicators on the screen. Fortunately, players learned these indicators during the game session, but they were not clear from the beginning, and it had a tremendous impact on their gaming performance. Usually, players started to investigate what these indicators meant once the moderator had noticed the possible problem and asked about it from the players. We were curious as to why playability experts did not report this problem, because it seemed like an obvious problem that should have been reported. The playability experts said that they did not consider the visualization of the indicators problematic, because it took only a short time for them to learn what they meant in the game.

From the gameplay, the players reported one particularly interesting problem that the playability experts did not report. Three players had problems in recognizing a new long-term goal, which was different than the long-term goal in the previous version of the game. Instead of having the game character avoid any obstacles in the game world, the players' main task is to maintain the energy level of the game character while performing short term goals (e.g. collecting items). The player can gain or lose energy by hitting obstacles in the game world. In addition, without new energy supplies, the energy level will drain slowly. In the previous version of the game, hitting an obstacle is lethal and the game character is self-sustained. The players did not notice this change and they were puzzled as to why the game character seemed to die without a reason. Other player-reported gameplay issues were mostly opinions rather than actual playability problems.

4.3 Expert Findings

Playability problems that were reported by the experts only were all related to the gameplay. The playability experts reported seven problems that were not discovered by the players (Table 2). Three of those problems were ranked as critical, but they were discovered from the levels that the players never played during the play session. However, it is very likely that these problems will materialize when the players reach those levels. The probability of encountering those problems is high, because they are all related to basic game mechanics. Two of these problems are related to collectable items in the game world. The game requires that items on some levels must be collected in a certain order, and if a player fails to do that, completing the level is impossible. These limitations will restrict the player's choice to complete a level and they may cause the game to stagnate. The third gameplay problem is related to substantially increased difficulty in the middle of the level. The game requires very

accurate navigation in order to collect items sequentially, and the pace of the game is temporarily increased at the same time. The actual problem is that the collectable items are possibly misplaced in the game world, and the input devices of the mobile phone do not enable the navigation accuracy needed on the level.

Other four playability problems discovered by the experts were related to the goals, rewards, and challenge of the game. The severity ranking of these problems was medium. The playability experts reported two problems related to the goals of the game. The game has a short term goal, which gives a special bonus if a player manages to collect all bonus items from the levels. However, these items are permanently attached to specific levels, and if a player starts a play session other than the first level, it is not possible to collect all bonus items and thus get the bonus.

The second problem with the goals was related to conflicting short term goals. The primary goal of the player is to avoid collision with obstacles which are not energy supplies, and to collect items in the game world. In the current version, this goal alone is a challenging task for the player. However, each level also includes a time limit, and the player fails if the time runs out before all items are collected. The reason why these two goals are conflicting is that controlling the game character to avoid collisions and collecting items is time consuming, and the player cannot really hurry up actions that need to be completed.

The playability problem related to the rewards of the game is that a good high score can only be achieved by playing levels in sequence and starting from level one. Completing more advanced levels does not give significantly more points than easier levels, and it is easier to complete easy levels than to try to play more difficult levels. The lack of proper rewards on the advanced levels can reduce the motivation of the player.

The playability problem related to challenge was found on the level that players never reached in the playtesting. The level contained an enemy which could destroy the game character instantly. This was unexpected feature, because the game character has an energy level and hitting an obstacle or an enemy will always decrease the energy level, but never drain it completely. Therefore, encountering such a strong enemy sounds like an unbalanced game feature and thus, it was reported.

4.4 Effectiveness of Evaluation Sessions

In playtesting, the participants' performance in terms of evaluation efficiency and how many playability problems they encountered was quite similar. On average, one session reported 20.17 (σ 2.23) playability problems and tendency was towards critical problems than minor problems (Table 2).

Table 2. Distribution of identified playability problems

	Mean	SD
Major	9.33	2.07
Medium	8.33	1.03
Minor	2.50	0.55

When evaluating the efficiency of the playability experts, we compared the average number of players who encountered a playability problem in playtesting which was also reported by the experts to those problems that were only reported by the players. The independent t-test indicates that there is a statistically significant difference in means, $t(43)=2.02$, $p<0.01$.

5. Discussion

Recent studies indicate that both the expert review and user testing have provided similar kinds of results [13], [27]. Our results also show that the expert review and playtesting identified playability problems quite consistently. There were altogether 53 playability problems identified from the game. 70% of them were related to user interface aspects such as visualization, feedback and controls, and 30% of the playability problems were related to gameplay issues. Distribution between the user interface and gameplay problems corresponds to the previous study [6]. One probable reason for this is that evaluating user interface issues is easier than finding gameplay, issues because problems in information visualization, feedback and navigation are apparent to both the players and the experts. An interesting observation from the results is that the expert review identified fewer problems than playtesting. Usually, Human-Computer Interaction (HCI) literature states that the expert review method tends to find more problems than user testing [13] (p.56), and quite many of them are false alarms that are not verified in the user testing [16], [11].

In this study, the expert review method was very well able to predict playability issues that cause problems for players. Both evaluation methods jointly reported 20 playability problems from the user interface of the game. Nine problems were ranked as critical problems, and 10 problems had medium severity. This indicates that both methods have found the most serious problems from the user interface. In addition, the playtesting identified 17 user interface problems. However, a detailed analysis of these problems indicated that usually only one out of six players encountered these problems, and more than half of them were ranked as minor in the severity classification.

For gameplay related issues, the diversity of the problems was greater. Both methods identified 5 common playability problems in the gameplay. Three of them were major problems, and 5-6 players encountered problems which were also reported by the experts. Two other gameplay problems were ranked with medium severity. In addition to jointly reported problems, the playability experts reported seven additional problems. During the evaluation session, the experts were able to progress further in the game, and they uncovered playability problems from levels that players never played in the playtesting. In addition, the experts uncovered playability problems related to goals and rewards that were not identified by the players. Reporting these kinds of problems can be regarded as specific benefit of the expert review method, because they require that the game is played repeatedly and the game mechanics are explored systematically.

The playtesting reported four playability problems that were not identified by the experts. One of them is a particularly interesting problem, because the players had difficulties understanding the long-term goal of the game as it conflicted with their previous experience of the earlier versions of the game. The game designers had dramatically altered the behavior of the game character and what the player must do to keep the character alive. The playability experts did not report such problem, because it was not directly related to the current game design. Of course, the experts should consider the previous experience and knowledge of the target users and how it might influence on gaming experience, but in this case, it was not a problem that everybody would face when playing the game. During the playtesting, half of the players reported this particular problem.

The results of our study indicate that expert review is an efficient evaluation method for evaluating the playability of mobile games, and the method can be used to complement results from playtesting. In our study, the evaluation procedure was similar in both cases, which will make the comparison of the results feasible. In both evaluations, the experts and players were able to play the game as they would normally do. We did not give any specific tasks that should be completed during the evaluation, but they were allowed to explore the game world as they would like to. Games, in general, are quite linear at the beginning and players are guided through the first missions or levels by the game design [1]. Therefore, the progress of the expert review will correspond to the progress of the players, and the experts can explore issues that are possibly problematic for players.

There are several advantages to using experts to evaluate game design instead of inviting players for playtesting sessions. Quite often, the playability experts have required domain expertise which is recommended by literature when recruiting evaluators [18]. The evaluators may belong to the target audience of the game, and playing games regularly will provide good baseline knowledge of different games. This is also a distinguishing aspect when compared to productivity software evaluations, where finding a domain experts can be much harder.

Another advantage of using playability experts is that although they are playing the game, their main focus is still in the evaluation, and they are analyzing their own behavior in order to identify possible playability problems in the game design. In this evaluation mode, they are more sensitive to recognizing problems and writing them down immediately. In playtesting, players are focusing on playing the game and more easily immersed by the gameplay. In this mode, they do not necessarily recognize problems they are facing, but instead they try to overcome them. This is a contradicting situation, since the main focus should be in identifying playability problems from the game, but if the game is good, players will be immersed by the game and it is difficult for them to express problems.

In our study, the players were frequently silent for a long period and playing intensively. Even though it is the moderator's responsibility to observe the behavior of the players, it is impossible to recognize all problems that the players facing in the game if they do not express them to the moderators. In addition, it is often very difficult for the moderator to start asking questions because it will break the immersion and disturb the player experience [26]. For many players, it seemed to be very difficult not to be too immersed and think aloud while playing the game. Therefore, we tried to direct our questions to proper moments (e.g. completing the level or restarting the level) that would disturb the game play the least. However, the problem with this approach was that the problematic moment had already passed, and the players were not always able to recall the problem that they were struggling with. Therefore, our observation supports Laitinen's findings [8].

The third advantage of using playability experts is that they will use the time reserved for evaluation efficiently, and from the evaluation point of view, they can evaluate the game more thoroughly than is possible in playtesting. In our study the expert review lasted three to four hours in total, and the experts managed to explore the game more thoroughly than the players in playtesting, although time for playing was approximately same. In our study, the playtesting sessions were limited to two hours. Completing all playtesting sessions took 12 hours. During playtesting, the players did not progress as far in the game as the playability experts. Especially if the players encountered a very challenging situation or a level, they spent lot of time trying to solve it. These challenges are part of the game, but if they do not contain any playability problems, they are only consuming valuable evaluation time. One solution to this problem is to provide cheat codes to the players, but this will probably distort evaluation results. In our study, the evaluators faced the same challenges as the players, but since they were also skilled players, they solved the challenges quicker and were then able to move ahead in the game. However, it should be noted that it is good to reserve more time for game evaluations than what is usually reserved for expert reviews of productivity software. Korhonen and Koivisto have previously noted that game evaluations will take longer, because the evaluators need to play the game and solve challenges that are included in the game design [6].

The fourth advantage of the expert review method is that the experts pay attention to issues that the players might ignore, but they are still important from an evaluation point of view. In our study, the experts reported a few problems related to goals and rewards which were not covered by the playtesting. The players were so keen on completing levels that they did not pay attention to rewards or short term goals at all. The players commented on rewards after the moderator asked about them, but responses indicated that they were not interesting at this point. The motivation for playing the game was targeted towards seeing new levels than analyzing the results. The playability experts, instead, explored how players will progress in the game and what is achieved by completing levels. These observations discovered new playability problems in the game design.

The expert review method does not make playtesting obsolete, but it should be seen as a complementary evaluation method which can provide useful information for game developers with less effort and pinpoint obvious problems before playtesting. In our study, the expert review identified playability problems consistently with the playtesting method. Moreover, the expert review was able to conduct a more thorough evaluation of the game than playtesting. However, the playtesting also reported some playability problems which were not identified by the playability experts. The most important finding was related to the conflict between the players' previous knowledge and assumptions and the current design of the long-term goal of the game. These kinds of findings are very difficult to achieve with the expert review method, because the evaluators' previous experience and knowledge influence what kinds of playability problems are identified. Another issue with the expert review is that it cannot describe the feelings and experiences that a game elicits from players, because the experts can only describe their own experiences. Therefore, playtesting is needed to explore the experiences of the game.

The limitations of the study are that we have compared the evaluation results of one mobile game. The results are, however, very encouraging and they provide initial findings that the expert review method provides useful data for game developers. Another limitation of the study is that we evaluated a game in which completing the first levels of the game is controlled very tightly by the game system. For other kinds of games, (e.g. MMORPGs or sandbox games) the evaluation procedure might need tighter

control as regards what aspects the evaluators should focus on during the evaluation.

In the future, we plan to continue these game evaluations and collect data from several game evaluations to validate the results presented in this study. In addition, we need to investigate the playtesting method and try to find ways of improving the effectiveness of the method in game evaluations and to overcome challenges that immersed players will set. Especially, we need to concentrate on how to identify playability problems in the game content, which is the most important part of the game.

6. Conclusion

In this paper, we have presented an evaluation study that compared the effectiveness of the expert review and playtesting methods in a mobile game evaluation. Playtesting is a commonly used evaluation method, but the expert review method is mainly ignored. In this study, we explored how the expert review method compares to playtesting in terms of efficiency in finding playability problems. The expert evaluation was conducted by two playability experts, and six players participated in the playtesting. The expert review used playability heuristics that are specifically designed for mobile game evaluations. Playability problems were analyzed and categorized based on their severity and heuristic that they violated. The results indicate that the expert review was accurately predicting playability problems that players faced when playing the game. The expert review discovered the most serious playability problems from the user interface that were also reported by the players. Playtesting reported many playability problems that are very detailed and specific to a certain player and which were not reported by the experts. Playability problems related to gameplay were much harder to discover in playtesting, and expert review found several serious problems that were not discovered by playtesting. We observed several benefits of the expert review method that will make it an attractive method for game evaluations. The length of the playtesting session will limit the scope of the evaluation, especially if the players encounter a tough challenge and cannot proceed in the game. Immersion into the game makes recognizing playability problems harder for players and expressing their difficulties to the moderator is often inadequate or even missing completely, because they are engaged in the gameplay. For these reasons, the expert review with playability heuristics can provide a cost efficient and fast method for evaluating the playability of a game.

7. ACKNOWLEDGMENTS

We would like to thank Elina Ollila for her help in arranging both evaluation sessions. In addition, we would like to thank Dr. Juha Arrasvuori and Dr. Matt Jones for their valuable comments on the previous versions of the manuscript.

8. REFERENCES

[1] Adams, E., Rollings, A. *Game Design and Development: Fundamentals of Game Design*. Prentice Hall, 2007

[2] Baauw, E., Bekker, M. M., Barendregt, W. A Structured Expert Evaluation Method for the Evaluation of Children's Computer Games. In proc INTERACT, (2005), 457-469

[3] Desurvire, H., Caplan, M., Toth, J. A. Using heuristics to evaluate the playability of games. In proc CHI Extended Abstracts, ACM Press (2004), 1509 - 1512

[4] Drachen, A., Canossa, A. Towards gameplay analysis via gameplay metrics. In proc MindTrek, ACM (2009), 202-209

[5] Dumas, J. S., Loring, B. *Moderating Usability Tests: Principles & Practices for Interacting*. Morgan-Kaufmann, 2008

[6] Federoff, M. A. Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games. Indiana University, 2002

[7] Fullerton, T., Swain, C., Hoffman, S. *Game Design Workshop: Designing, Prototyping, and Playtesting Games*. CMP Books, 2004

[8] Gray, W. D., Salzman, M. C. Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. *13,* 3 (1998), 203-261

[9] Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B., Pagulayan, R. J., Wixon, D. Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. In proc CHI, ACM (2008), 443-452

[10] Korhonen, H., Koivisto, E. M. I. Playability Heuristics for Mobile Game. In proc MobileHCI, ACM Press (2006), 9-16

[11] Korhonen, H., Paavilainen, J., Saarenpää, H. Expert Review Method in Game Evaluations – Comparison of Two Playability Heuristic Sets. In proc MindTrek, ACM Press (2009), 74-81

[12] Laitinen, S. Do usability expert evaluation and test provide novel and useful data for game development? *JUS 1,* 2 (2006), 64-75

[13] Molich, R., Dumas, J. S. Comparative Usability Evaluation (CUE-4). *BIT 27,* 3 (2008), 263-281

[14] Molich, R., Kaasgaard, K., Karyukin, B. Comparative Usability Evaluation. *BIT 23,* 1 (2004), 65-74

[15] Mulligan, J., Patrovsky, B. *Developing Online Games: An Insider's Guide*. New Riders, 2003

[16] Nielsen, J. Finding usability problems through heuristic evaluation. In proc CHI, ACM Press (1992), 373-380

[17] Nielsen, J. Heuristic Evaluation. in Usability Inspection Methods, Nielsen, J., Mack, R. (Eds) Wiley & Sons (1994), 25-62

[18] Nielsen, J. *Usability Engineering*. Academic Press, 1994

[19] Nielsen, J. Usability inspection methods. In proc CHI Extended Abstracts, ACM (1994),

[20] Nielsen, J., Molich, R. Heuristic evaluation of user interfaces. In proc CHI, ACM (1990), 249-256

[21] Nielsen, J., Phillips, V. L. Estimating the relative usability of two interfaces: heuristic, formal, and empirical methods compared. In proc CHI, ACM (1993), 214-221

[22] Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., Fuller, T. User-centered Design in Games. in Handbook for Human-Computer Interaction in Interactive Systems., Jacko, J., Sears, A. (Eds) Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (2003), 883-906

[23] Pinelle, S., Wong, N., Stach, T. Heuristic evaluation for games: usability principles for video game design. In proc CHI, ACM Press (2008), 1453-1462

[24] Rouse, R. *Game Design: Theory and Practice*. Wordware Publishing, 2001

[25] Rubin, J. *Handbook of Usability Testing: How to plan, design and conduct effective tests*. John Wiley & Sons, 1994

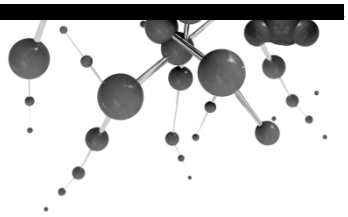
[26] Schell, J. *The Art of Game Design*. Morgan Kaufmann, 2008

[27] Tan, W.-s., Liu, D., Bishu, R. Web evaluation: Heuristic evaluation vs. User Testing. *39,* 4 (2009), 621-627

9. Appendix A

Game Usability Heuristics	
GU1	**Audio-visual representation supports the game**
	The game graphics should support gameplay and story and be informative for the player. In addition, the graphical look and feel should be consistent throughout the game. Audio can be used to evoke emotions and increase immersion. A good sound environment in the game supports a positive gaming experience. The graphics or audio should not prevent the player from performing actions or make them unnecessarily difficult.
GU2	**Screen layout is efficient and visually pleasing**
	The layout should present all necessary information for the player, but on the other hand, if the screen is filled with all kinds of information, it starts to look crowded. It is important that the player finds the navigation controls and they should not be mixed with the information that needs to be visible on the screen.
GU3	**Device UI and game UI are used for their own purposes**
	It should always be noticeable whether the player is dealing with the game user interface or device functions. The game interface should not use the device's user interface widgets in the game interface, because it breaks the immersion. The most impressive immersion is achieved when the game uses full-screen mode hiding other features.
GU4	**Indicators are visible**
	The player should see the information such as the current state of the game and status of the game character that is required for being able to play the game. Information that is frequently needed should be visible for the player all the time — if possible.
GU5	**The player understands the terminology**
	The terminology that is used in the game should be understandable and not misleading or unfamiliar for the players. Technical jargon should be avoided. For instance, terminology that is related to the game concept or features that the game needs from the device should be translated into more understandable language.
GU6	**Navigation is consistent, logical, and minimalist**
	Navigation consists of the game menu and the game world. The game menu consists of settings and selections for the desired game session. Different functions should be organized reasonably and possibly on different screens. However, long navigation paths should be avoided. Short navigation paths provide more clarity and are easier to remember. In the main game menu, the player should be able to start a game and have access to other important game features. In the game world, navigation should be intuitive and natural. Regardless of the complexity of the game world, players should be able to navigate there smoothly. With a proper set of control keys, navigation can be very intuitive and almost invisible.
GU7	**Control keys are consistent and follow standard conventions**
	Using common conventions in control keys reduces the time that is needed to learn to play the game since the player can use his or her knowledge from other games. Game devices usually have specific keys for certain actions and every game should follow them.
GU8	**Game controls are convenient and flexible**
	Novice players usually need only a subset of the controls when they start playing the game. On the other hand, veteran players often need shortcuts and more advanced commands. It should be possible to customize the game controls or use shortcuts or macros. However, using shortcuts should not provide a major edge in a competitive player vs. player game. The configurability and amount of controls needed to play the game should be kept at the minimum, but they need to be sufficient. In addition, the controls should be designed according to the device's capabilities.
GU9	**The game gives feedback on the player's actions**
	A good user interface has a low response time on the player's actions. An action can be either a single key press or a more complicated input sequence. The player should notice immediately that the game has recognized the action by providing feedback. The most common way of providing feedback is to present it graphically. Other alternatives are to use audio or tactile feedback. Providing only auditory feedback is not acceptable since a player may be playing the game without sounds. Although the game needs to respond immediately to the player's actions, the consequences of the action can be shown to the player later. If an action cannot be performed immediately, the game should notify the player about the delay.
GU10	**The player cannot make irreversible errors**
	The game UI should confirm actions that can cause serious and irreversible damage, which affects the player's ability to play the game. Such errors are typically related to the game character or player's progress in the game. When mistakes happen, it is helpful to enable recovery.
GU11	**The player does not have to memorize things unnecessarily**
	The game should not stress the user's memory unnecessarily, unless it is part of the gameplay.
GU12	**The game contains help**
	The players do not often read manuals. Instead, the game should teach the player what he or she needs to know to start playing the game. This can be done through a tutorial mode at the beginning of the game. The tutorial mode should be divided into chapters that teach a couple of things at the beginning. Ideally, the tutorial could be embedded completely in the game so that help would be provided every time when it is really needed. Help is also often needed in error situations. If the game provides useful error messages, the player can understand better what caused the problem.

Gameplay Heuristics	
GP1	**The game provides clear goals or supports player-created goals**
	The players should be able to understand goals that exist in the game. The goals can be either set by the game or created by the players. The game should contain both short-term and long-term goals. Short-term goals provide repeated opportunities for reinforcement and keep players motivated to play the game. Long-term goals are usually more difficult to achieve and they can consist of several short term goals.
GP2	**The player sees the progress in the game and can compare the results**
	The players should have enough information so that they can see their progress towards the goals in the game. The progress can be shown to the player explicitly, for instance with numbers, or implicitly, for instance, by changing the behavior of non-payer characters or the game world. The players feel more motivated if they can compare themselves with the other players or the previous achievements. Traditionally, this has been done with high-score lists, rankings, character levels, or different titles.
GP3	**The players are rewarded and rewards are meaningful**
	The players should receive a meaningful reward as they progress in the game. In addition, the reward should be adjusted to the challenge that the player had to face in order to get it. The rewards schedule should be varying and frequent, but still unpredictable.
GP4	**The player is in control**
	The game should provide at least an illusion that the player is in control of what is happening in the game world. The players should be able to decide on actions they want to take and these actions should have an influence on the game world.
GP5	**Challenge, strategy, and pace are in balance**
	The game should be designed so that the challenge is comparable to player's current skills, then the players do not feel frustrated or bored with the game. In single-player games, the player can often choose the difficulty level and thus affect the challenge. The players learn new strategies as they play the game. There should not be dominating strategies for any part of the game. The pace should be adjusted to the game style and it can be intensive or deliberate. The game should allow the player to take a deep breath once in a while during the play sessions.
GP6	**The first-time experience is encouraging**
	The first impression of the game is formed within a few minutes and it is very difficult to change. The players should feel that they have learned the basics and have accomplished something. The first play session should make the player desire for the next play session.
GP7	**The game story supports the gameplay and is meaningful**
	Even though the story plays an important role in many games, it should not dominate the gameplay. Some games do not even have or need a game story. If the game has a story, it should fit the other elements in the game and sound plausible to the player. The dialogue with non-player characters (NPC) should be meaningful and interesting for the player.
GP8	**There are no repetitive or boring tasks**
	The game should not require repetition of tasks without changing any conditions. Often, this repetition happens when the player needs to reach a certain goal before the game becomes interesting or challenging. However, during the training phase (tutorials), it is useful to repeat certain tasks so that the player learns and practices for example how the character is controlled in the game.
GP9	**The players can express themselves**
	The players should be able express themselves by, for instance, customizing their characters, acting in a certain way, or modifying the game world. Allowing the players to customize and personalize their game characters makes it more probable that they feel attachment to a game.
GP10	**The game supports different playing styles**
	The players can vary a lot in terms of both experience and preferred play styles. There are also different playing styles that should be supported at least in the more complex games. The player types are defined based on how the players prefer to interact with the game world or with the other players, Four common player types are: A) Achievers, who like to compete with the game mechanics. B) Explorers, who wish to explore different aspects of the game. C) Socializers, who prefer to socialize with other players D) Killers, who enjoy dominating other players.
GP11	**The game does not stagnate**
	The players should always feel that it is possible to reach the goals and the game progresses. The game should recognize immediately when the game is over and inform the players. Ending of the play session should be clearly indicated and restarting the game should be possible.
GP12	**The game is consistent**
	The game world and actions should be consistent and logical for the player. If something works in the beginning, the player assumes that it also works later on. Correspondingly, if the player is able to perform a certain action in the game world or for a game item, the player assumes that similar kind of action is possible for other similar objects or in the similar situation as well. Moreover, if the game world resembles the real world, the player assumes that the same principles also work in the game world. The game should not contain invisible walls.
GP13	**The game uses orthogonal unit differentiation**
	Each game item should have a purpose in the game world and it should be notably different to other similar game items. In addition, if the player needs to select character classes or roles in the game, they should be functionally different.
GP14	**The player does not lose any hard-won possessions**
	The game should maintain possessions that the player has earned while playing the game and the player cannot lose them accidentally. However, in some cases the game can provide very high risks and the player can stake valuable game items which can be lost during the gameplay.



Paper VI

Korhonen, H. (2011). The Explanatory Power of Playability Heuristics. In *Proceedings of the Advances in Computer Entertainment Technology (ACE'11, Lisbon, Portugal)*, Article No. 40. New York, NY, USA: ACM, doi: 10.1145/2071423.2071473

Accepted for Publication in ACM Computers in Entertainment Journal

The Explanatory Power of Playability Heuristics

Hannu Korhonen
Nokia Research
P.O. Box 1000
00045 Nokia Group, Finland
hannu.j.korhonen@nokia.com

ABSTRACT
Research of playability heuristics to be used with usability inspection methods in videogame evaluations has been active in recent years, and there are multiple playability heuristic sets available. However, they differ quite a lot from each other, and it is still unknown how well they support inspectors and help to describe the identified playability problems in a game. In this paper, we present a study in which 36 novice inspectors were divided into two identical groups, and they evaluated a mobile game using one of the two playability heuristic sets. In the second task, the inspectors analyzed playability problems that were collected from the game, and assigned heuristics to describe the problems. The results show that a playability heuristic set needs to cover the main aspects of playability in order to be effectively used during the evaluation. The results of this work will help in developing a playability heuristic set that makes expert review an effective method for evaluating videogames.

Categories and Subject Descriptors
H.5.2 [**Information Systems**]: Information Interfaces and Presentation – *User Interfaces, Evaluation/Methodology.*

General Terms
Experimentation, Human Factors.

Keywords
Playability, Heuristic, Expert Review, Game Evaluation, Domain Specific Heuristics, Comparison Study.

1. INTRODUCTION
Playability is a crucial factor when determining successful videogames which provide an engaging gaming experience for players. At least two aspects influence the playability of the game. The user interface should be intuitive and unobtrusive, so that it does not disturb gaming, and players can concentrate on playing the game. Secondly, gameplay should be understandable, balanced, sufficiently challenging, and engaging [1]. In addition to this, there can be other aspects such as social interaction in multiplayer games which should be taken into account when evaluating playability.

Game designers have relied mainly on player feedback when evaluating videogames [2], [3], [4]. Focus groups are used to collect ideas and evaluate game concepts in the early stages of game development. At the later stages, playtesting is used to uncover design problems in a game and collect feedback from players. However, for conducting evaluations effectively, there should be a complete working game prototype available that the players can play during the evaluation.

Analytical inspection methods, on the other hand, can provide a cost-efficient way to evaluate game design with partial or even non-functional prototypes, and to collect useful feedback for designers prior to involving players in evaluations. Expert review is a commonly used method for evaluating the usability of traditional software. The method is originally developed by Nielsen and Molich in 1990 [5]. In the evaluation, usability specialists go through the design and identify issues that are likely to cause problems for end-users. The expert review method does not provide information on how players will experience the game, but the specialists can predict problems that the players might face, which in turn decreases the enjoyment that the game is designed for. Even persons who are not necessarily usability specialists are able to conduct an evaluation after a short training [5]. However, the method has some challenges when it is applied to game evaluations. The traditional usability heuristics are not applicable to game evaluations, because they contravene in achieving a good game experience [6] or they do not cover all important aspects of the games [7]. Nielsen has stated that there should be domain-specific heuristics for specific products [8], such as videogames, to complement the general heuristics.

Evaluating videogames is one of the few areas in which the expert review method has not been applied commonly, but the interest in using the method in game evaluations has increased, and there are several domain-specific heuristic sets available that are targeted towards evaluating the playability of videogames (e.g. [9], [7], [10], [11]). The development of playability heuristics, however, is still in its infancy, and the scope and the content of the heuristic sets seem to differ quite substantially from each other. Although multiple playability heuristic sets exist, there is no exact knowledge on how inspectors are using those heuristics during the game evaluations, and whether they actually help in identifying playability problems. One characteristic of the expert review method is that inspectors use their own expertise and previous knowledge about game design to identify playability problems during the evaluation. The role of the heuristics is to support the inspectors and provide help in paying attention to the most critical aspects that will influence playability. Knowledge of the usefulness of the playability heuristics would contribute in developing an optimal heuristic set that comprises the most important or commonly occurring playability problems in videogames. Currently, only limited knowledge exists on how different playability heuristic sets compare in a videogame evaluation [1], [12], but the actual use of the heuristics has not been investigated.

In this study, we explore how novice inspectors use playability heuristic sets in a game evaluation, and explore in detail what playability heuristics are referred to when assigning violated heuristics for playability problems.

Full presentation, ACE'2011 - Lisbon, Portugal

2. RELATED WORK

The development of usability heuristics or domain-specific heuristics includes a validation step, in which heuristics are evaluated for assessing their relevance to describe usability problems. The common approach is to compare each heuristic to playability problems and rate how well the heuristic covers the description of the problem [13], [14], [15]. In the development of the usability heuristics, Nielsen used 249 usability problems collected from the earlier studies to rate 101 usability heuristics [13]. As a result, Nielsen defined the current set of usability heuristics which are widely used with the expert review method.

Somervell and McCrickard performed a comparison of heuristic sets by providing a list of usability problems identified from a target system, and asked inspectors to rate the applicability of the heuristics to each of the problems [15]. The results indicate that more specific heuristics held the best scores [15]. The authors concluded that by providing a set of usability problems to the evaluators, it is possible to determine more accurately the applicability of the heuristic set to the problem set [15].

Assessing the relevance of the usability heuristics to each problem separately is made possible by the fact that both the heuristics and the problems concentrate on the same aspect of a product. This approach is not feasible when developing heuristics to evaluate games, because playability is dependent on two substantially different aspects: user interface and gameplay. It is obvious that gameplay heuristics are not good at describing user interface issues, and vice versa. Therefore, the assessment of the playability heuristics should be covered differently, and in the experiment, we have focused on the actual use of the heuristics in a game evaluation. The validation of playability heuristics is still in progress, and there are only a few studies in which playability heuristic sets are compared to other playability heuristic sets or usability heuristics.

Jegers has compared playability heuristics and Nielsen's usability heuristics in a pervasive game evaluation. 16 playability problems were analyzed, and the results indicate that some problems could be described with usability heuristics, but playability heuristics were better in describing problems [12]. Mankoff et al. concluded that because Nielsen's heuristics are quite general, the inspectors are usually able to assign those heuristics to describe problems even when there is not an obvious fit [14].

Korhonen et al. compared two playability heuristic sets in a mobile game evaluation. The results indicate that both heuristic sets need to be improved before they can be applied by the practitioners [1]. There were some problems especially in the clarity and understandability of the heuristics. The problems with the heuristics were visible in inspectors' ability to assign violated heuristics to describe a playability problem. In addition, different playability heuristics [9], [10], [11] have been validated independently, but these studies do not provide data on how well the heuristics supported the inspectors' task during the evaluation.

3. EXPERIMENT

Our aim is to investigate how playability heuristics sets were used in identifying and analyzing playability problems in a mobile game evaluation, and especially, which heuristics were referred to when reporting playability problems. We arranged a study in which two similar groups of novice inspectors evaluated a mobile game. The use of the playability heuristics was investigated in two tasks. First, the inspectors evaluated the game and reported playability problems that they had identified in the game. In the second task, the inspectors analyzed a predefined list of playability problems from the game and assigned violated heuristics to each playability problem. The predefined playability problem list was constructed by a playability expert who evaluated the game prior to the novice inspectors.

3.1 Inspectors

The experiment started with inspector group formation. We recruited 36 students from the local university to participate in a game evaluation. The students were attending a basic usability evaluation course during the semester, and they had basic knowledge of the expert review method, as it was one of the methods of the course. In addition, they had conducted a heuristic evaluation of a web application before participating in our experiment. Some students had conducted usability evaluations using user testing and/or heuristic evaluation as part of some other course prior to attending the course at the time of the experiment. All participants can still be considered novice inspectors, because their evaluation experience is limited to evaluations included in the university courses.

The objective was to form two similar inspector groups for the experiment. We collected background information, and the group formation was done based on demographics, gaming activity, and evaluation experience (Table 1).

Table 1 Basic statistics of the inspector groups

Group Name	H06	H08
Males	14	15
Females	4	3
Mean Age (in years)	23,29	24,28
Major Subject	Computer Science Interactive Techology Statistics	Computer Science Interactive Technology
Evaluation Experience (# of students)		
Heuristic Evaluation	7	5
User Testing	4	6

The gaming activity of the inspectors varied, and they could be roughly divided into two categories. About half of the inspectors in both groups played less than two hours per week or they played videogames infrequently. The other half played more, and in both groups, there were a few inspectors who played up to 40 hours per week. There were no particularly popular game genres, and the inspectors played different kinds of games. The most popular gaming platforms were PC, Game console (Xbox, PlayStation or Wii), Handheld device (PSP, Nintendo DS), and mobile phones.

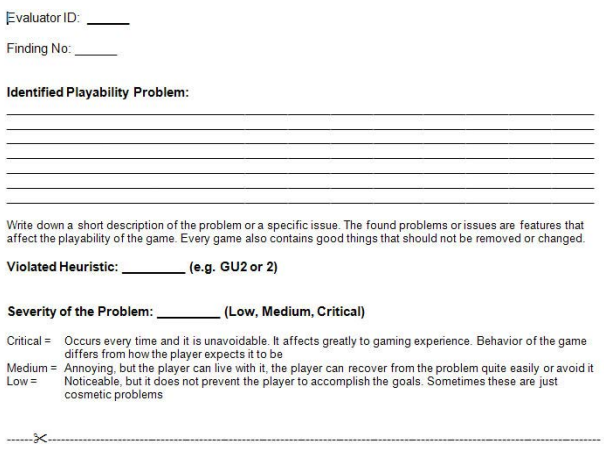
3.2 Procedure

After the group formation, the inspectors received the heuristics and instructions concerning how to prepare for the evaluation one week before the evaluation sessions. The inspectors were instructed to study the playability heuristic sets in their own time. Moreover, it was stated that the inspectors did not have to memorize the heuristics, but that they would have the heuristic set with them during the evaluation. 30% of the inspectors had read the heuristic set once and others had read it 2 to 3 times. The group assignment and the game were not disclosed to the inspectors at this point.

We arranged three evaluation sessions during two consecutive days, and the inspectors were free to choose which session they attended. Each inspector could only participate in one session. At the beginning of the session, the inspectors received a mobile phone with the game installed on it, a headset, and an envelope which contained a Non-Disclosure Agreement (NDA) form, the

heuristic set, a pile of problem report forms, and a post-test questionnaire. The NDA form was needed because the game was an unpublished commercial game under development.

Before the inspectors started the evaluation, we briefly went through how to report playability problems, what to do if the game or phone crashed, and what the game is about. In addition, we introduced the basic controls (navigation, volume and special function keys) of the mobile phone, because the inspectors were not familiar with the phone models. We also reminded the inspectors that the game is in alpha testing phase, meaning that some features might be missing, and that there are obvious bugs which can be ignored during the evaluation. Detailed instructions on how to play the game were not given to the inspectors. The inspectors were advised to write down issues that they observed using a problem report (Figure 1), and to report one problem per report. The problem criterion was "*Is there a playability issue that disturbs your gaming experience in the game?*" The inspector described a problem using his or her own words, assigned a violated heuristic, and defined the severity using three severity values (critical, medium, low). The inspectors were not asked to provide a recommendation on how to fix the problem.


Evaluator ID: _____

Finding No: _____

Identified Playability Problem:

Write down a short description of the problem or a specific issue. The found problems or issues are features that affect the playability of the game. Every game also contains good things that should not be removed or changed.

Violated Heuristic: _____ (e.g. GU2 or 2)

Severity of the Problem: _____ (Low, Medium, Critical)

Critical = Occurs every time and it is unavoidable. It affects greatly to gaming experience. Behavior of the game differs from how the player expects it to be
Medium = Annoying, but the player can live with it, the player can recover from the problem quite easily or avoid it
Low = Noticeable, but it does not prevent the player to accomplish the goals. Sometimes these are just cosmetic problems

Figure 1 Playability Problem Report

The inspectors started the game from scratch, meaning that their first task was to create a game avatar. Then they had to move to song selection, and finally to control the dancing of the avatar. We did not specify any particular objectives or scenarios that the inspectors should try to achieve or follow; they were free to play the game as they liked. This approach provides a realistic context for the inspectors, and it resembles the situation that players will face when they get a new game for their device. In addition, in game evaluations, defining tasks and scenarios beforehand is not usually necessary, because the game should introduce goals for a player and guide the first moments of the play session. The same method is also used in game evaluations at the Microsoft Games User-Testing Group [16]. The session lasted for two hours, and the inspectors conducted the evaluation independently. We did not allow inspectors to discuss their observations during the evaluation session, as we wanted to collect the original list of problems that each inspector identified. The inspectors were also advised to use the headset during the evaluation in order to minimize cacophony in the classroom.

At the end of the evaluation session, the inspectors filled in a post-test questionnaire in which we asked how the heuristic set had worked for them and how the inspector felt about using the heuristic set during the evaluation. The questionnaire contained closed questions offering a number of defined responses (using the 7-point Likert scale). We also asked how the inspectors had prepared for the session. The post-test questionnaire, the problem reports, other forms, and the mobile phone were returned in the closed envelope that they received when they arrived at the evaluation session. The inspectors were asked not to discuss the evaluation session with other course participants until the whole experiment was finished. The inspectors were rewarded with extra course credit points for their participation in the game evaluation.

In the second task of the experiment, the inspectors had to analyze a predefined list of playability problems that appeared in the game. The inspectors received a playability problem list and the post-test questionnaire by email after the final evaluation session. The problem list was created by an experienced playability expert (with expertise of videogames and the method). Previous studies have concluded that experienced inspectors, or experts, produce more complete sets of usability problems than novice inspectors [8], [17]. The previous results from the comparison study of mobile game evaluation indicate that a playability expert can very accurately identify major and medium-level playability problems that players will face when playing the game [18]. Thus it can be assumed that an experienced playability expert can provide a problem list that can be used for validating the evaluation results of the novice inspectors. The inspectors received movie tickets as rewards after completing the second task.

The playability problem list was created prior to the evaluation sessions, and the expert followed same procedure and used the same equipment as novice inspectors would use in the experiment. Each problem was described in a couple of sentences. The expert was very familiar with both heuristic sets that were selected for the study. For each problem, the expert assigned a violated heuristic from both sets, if there was a suitable heuristic in the set. The list contained 44 playability problems from the game.

After all evaluation sessions were over, we analyzed the playability problem reports. The analysis was based on the problem description that an inspector had written. In addition, we recorded the severity rating and violated heuristics that the inspector had defined for problems. If the problem description contained multiple playability problems, those reports were split and the problems were analyzed separately. The problem reports were explored to find out in which game feature or part of the game the problem occurred. The problems were categorized into eight groups: *Game Menu*, *Character Creation*, *Music Selection*, *Dancing*, *Photo Gallery*, *Mobility*, *Device*, and *Network*. Furthermore, each problem report was analyzed to determine whether the playability problem described in the report was related to user interface, gameplay, mobility, or device.

The analysis of the predefined problem list included recording of the violated playability heuristics, and comparing them to the violated heuristic assigned by the playability expert. Statistical methods were used to analyze the results of the post-test questionnaires from both tasks.

3.3 Apparatus and the Game

The inspectors mainly used Nokia 6210 Navigator mobile phones in the experiment (Figure 2). Another phone model, Nokia N81, was available as a backup in a case of phone malfunction, or if there were not enough phones available in the evaluation session. The phones are very similar in terms of the features in their technical specifications and industrial design that are relevant for playing games. During the experiment, 88% of inspectors used the Nokia 6210 Navigator model. Four inspectors (two in both groups) conducted the evaluation by using Nokia N81 phones.

Figure 2 Mobile phones (Nokia 6210 Navigator on the left)

The game was a new kind of music game for mobile phones. It uses the music collection stored on the phone as game content, and the player performs dance movements according to the rhythm and tempo of a song. Other major features in the game are creating and customizing a game character and managing songs that are played in the dance section. For evaluation purposes, we uploaded a music collection of 32 songs representing the dance, pop, hip hop and rock music genres to the phones.

A player controls a 3D character in the game (see Figure 2) by pressing the rocker key (4-way navigation on one physical key) and trying to make different dance combos while maintaining the rhythm of the song. In addition, the player is able to customize the character with new clothes that are unlocked as rewards. The player's performance is judged based on the complexity of the movements and the accuracy of the dancing, with the score presented as stars.

The game was designed and produced by a 3rd party game development company, and none of the participants, experiment organizers, or the playability expert had seen the game beforehand.

3.4 Selected Playability Heuristic Sets

The primary goal of the study is to provide an answer to the question of how well the playability heuristic set supports an inspector in the game evaluation. For this study, we selected two heuristic sets that are quite different in many respects and, therefore, some differences in how they assist inspectors in completing an evaluation can be expected. Both heuristic sets are general purpose heuristics, meaning that they are applicable for evaluating all kinds of videogames, regardless of their platform. For purpose of the experiment, both heuristic sets were presented in a similar way, including a heading followed by a short description of the heuristic.

3.4.1 Playability Heuristics for Mobile Games

The heuristic set developed by Korhonen and Koivisto contains 29 heuristics organized into three modules [10]. Each module can be included or excluded depending on the needs of the evaluation. Two core modules, *Gameplay* and *Game Usability*, are common for all games. The *Game Usability* module covers game controls and the interface through which the player interacts with the game. The *Gameplay* module contains heuristics that cover the game mechanics and other issues in the game content. The *Mobility* module contains heuristics that are specific for mobile games. The heuristics were developed based on a review of literature; the initial heuristics were used in game evaluations, and they were reviewed by game researchers and game designers [10]. The playability heuristics have been developed further in game development projects.

For this study, we modified the heuristic set based on the recommendations from the earlier study [18]. The group of inspectors who used this heuristic set during the evaluation is referred to as H06 in the results section.

3.4.2 Game Usability Heuristics

The heuristic set developed by Pinelle et al. contains 10 game usability heuristics that are based on analyzing PC game reviews from a popular gaming web site. The analysis included 108 games from six major genres, and it identified common problem categories which were used to develop these heuristics [11]. The heuristics are developed specifically to evaluate game-specific usability issues related to learning, controlling and understanding the game [11]. The set is missing heuristics that cover gameplay issues. The group of inspectors who used this heuristic set during the evaluation is referred to as H08 in the results section.

4. RESULTS

In this section, we describe how novice inspectors used heuristic sets during the game evaluation and the problem analysis tasks.

4.1 Heuristic Sets in Game Evaluation

The overall reporting activity of the inspector groups was quite similar. The inspectors reported 480 playability problems in total (Table 2). Group H06 reported 250 problems and group H08 reported 230 problems. The inspectors' low experience level in conducting the evaluation and reporting findings can be seen in the relatively high number of problem reports that contained minor bugs and other known technical issues. These issues were due to unfinished implementation, and the inspectors could have ignored them in the evaluation because they were already discovered in quality assurance testing and it was mentioned in the introduction session. In group H06, 19% of the problems were non-playability issues. In group H08, this number was substantially higher, and 26% of the reported problems were not referring to playability issues. After removing bug reports, duplicates, and other irrelevant reports (22% of total reports), there were 374 reports related to the playability of the game.

Table 2 Reporting activity by groups

	H06	*H08*
Total Reports	250	230
Playability problems	203 81%	171 74%
Non-playability issues	47 19%	59 26%
Mean Reports per Evaluator	11.28	9.50
Standard Deviation	3.51	3.94

We categorized these 374 playability problems and identified 64 unique playability problems. 45 playability problems were found in the user interface, 14 playability problems were gameplay problems, and 5 problems were related to mobility and device. 42 out of 64 playability problems were identified by inspectors from both groups. There were 17 playability problems that at least nine inspectors (25% of the inspectors) reported.

In this paper, we focus on these 17 playability problems in more detail to see how the inspectors have assigned violated heuristics when reporting problems, and to compare them to heuristics that a playability expert has assigned. The playability expert assigned a violated heuristic for 17 problems by using both playability heuristic sets, and it was done without knowing how the inspectors had assigned them. As the playability heuristic set used by group H08 is more limited than the set used by group H06, the playability expert could not assign a violated heuristic from this set for nine playability problems. The playability problems were

related to navigation and gameplay features, such as goal setting in the game. Table 3 shows a summary of the comparison between novice inspectors and the playability expert.

In general, the inspectors assigned violated heuristics accurately. In both groups, the inspector assigned on average 1.08 heuristics, and in seven cases, more than one heuristic was assigned. The playability expert assigned one violated heuristic for each playability problem. It is not forbidden to assign multiple violated heuristics, but if a problem report describes an issue with enough detail, usually one violated heuristic is sufficient.

The correspondence between violated heuristics assigned by the inspectors and the playability expert is not very high. In group H06, the inspectors achieved 35% accuracy. In group H08, the result is slightly lower, and 33% of the inspectors had selected the same violated heuristic as the playability expert.

In group H06, there were four playability problems in which none of the novice inspectors had selected the same heuristic as the expert. These problems were related to both user interface and gameplay. Correspondingly, in group H08, there were also four playability problems in which the playability expert had assigned a heuristic that none of the novice inspectors had selected. There was one playability problem ('Navigation – Missing Key Labels') for which inspectors in both groups assigned a different violated heuristic than the expert.

Another observation from the violated heuristics is that the inspectors were not very consistent in assigning violated heuristics, and they had assigned numerous heuristics to describe the problem. In group H06, the median value of the violated heuristics is four (maximum is seven and minimum is two). In group H08, the median value is three. Only in one case had all the inspectors in group H08 selected the same violated heuristic. This would indicate that the inspectors had not really been careful when assigning violated heuristics, because there are also very obvious cases in which assigning a violated heuristic would have been a straightforward task. For example, the most commonly identified playability problem in both groups was related to insufficient instructions for a player to start playing the game. In both playability heuristic sets, there is a heuristic that specifically points out the importance of an in-game help or tutorial; still only half of the inspectors (45% in group H06 and 56 % in group H08) had assigned this particular heuristic to describe the problem. Others had interpreted that the problem means lack of proper goals and missing indicators or misleading terminology in the user interface. There were also some inspectors who had not assigned any violated heuristic for the problem.

The inspectors had also identified another problem stating that the game is missing a tutorial that would teach the player basic interaction with a game avatar. In group H08, all inspectors had assigned the correct heuristic, but in group H06, only 64% of the inspectors had found the right heuristic. However, it should be mentioned that three inspectors in group H06 had assigned a heuristic that could be considered to describe the problem, because it is related to the initial impressions that a game gives the player, and how the first play session influences the player's willingness to continue playing the game. The playability expert had interpreted that the problem is related to the user interface and not gameplay aspects, which the heuristic is referring to.

The most accurate assignments of the violated heuristics in both groups were related to a problem that describes problems of challenge in the game. In group H06, 80% of the inspectors had assigned the same heuristic as the expert. In group H08, all inspectors had assigned a heuristic related to challenge in the game. In addition, one inspector had assigned a secondary heuristic for the problem.

Table 3 Summary of violated heuristic assignment during the evaluation

		H06				H08			
Category	**Playability Problem**	Inspectors	Violated Heuristics	Assigned Heuristics	Agree with Expert	Inspectors	Assigned Heuristics	Per Inspector	Agree with Expert
User Interface	Help - Insufficient Instructions	11	7	1,27	45 %	9	3	1,00	56 %
Gameplay	Goal - Goals are Missing	12	6	1,08	67 %	5	3	1,00	40 %
User Interface	Navigation - Start Game	6	2	1,00	83 %	9	5	1,11	11 %
Gameplay	Help - No Tutorial	11	3	1,00	64 %	4	1	1,00	100 %
Gameplay	Dance - Movements Arbitrary	6	5	1,00	33 %	7	6	1,29	29 %
User Interface	Feedback - Dance Visualization	7	4	1,00	14 %	5	5	1,20	40 %
User Interface	Navigation - Menu Structure	7	4	1,00	29 %	5	3	1,00	0 %
User Interface	Navigation - Confirm Selection	8	4	1,00	63 %	4	3	1,00	0 %
User Interface	Feedback - Flying Stars Confusing	6	4	1,00	33 %	5	4	1,20	40 %
Gameplay	Dance - Challenge Mode Confusing	5	5	1,00	0 %	6	6	1,17	17 %
User Interface	Navigation - Missing Key Labels	4	3	1,25	0 %	6	2	1,00	0 %
Gameplay	Challenge - Challenge Level	5	2	1,00	80 %	5	2	1,20	100 %
User Interface	Feedback - Song Duration	5	3	1,20	40 %	5	4	1,00	40 %
Gameplay	Points - Collection Unknown	7	5	1,14	0 %	3	3	1,00	33 %
User Interface	Navigation - Pause Menu Structure	7	6	1,14	29 %	2	2	1,00	0 %
User Interface	Navigation - Missing Commands	3	4	1,33	0 %	6	4	1,17	17 %
User Interface	Navigation - Default Menu Item	6	4	1,00	17 %	3	3	1,00	33 %
			4	**1,08**	**35 %**		**3**	**1,08**	**33 %**

4.2 Heuristic Sets in Problem Analysis

After the game evaluation sessions, we sent a list of playability problems that a playability expert had identified in the game to the novice inspectors. Their task was to analyze the problems and assign a violated heuristic to them. The list contained 44 playability problems from three main features of the game: *Character Creation*, *Music Selection* and *Dancing*. Each game feature contained problems which were related to the user interface and the gameplay. In addition, the music selection feature contained one problem which was related to the mobility aspect of the game. Each playability problem was described in a few sentences, and there was no indication of whether it was related to the user interface or the gameplay. The list and the post-test questionnaire were sent to the inspectors via email, and the inspectors were asked to return the assignment within one week.

The response rate in group H06 was 44%, and in group H08 it was 61%. The inspectors used the same playability heuristic set as they had used during the game evaluation. The inspectors were asked to assign as many violated heuristics to the problem as they saw fit, or leave the violated heuristic empty, if they could not assign an appropriate heuristic for a problem. Table 4 shows a summary of how novice inspectors assigned violated heuristics to the playability problems, and how they compare to violated heuristics assigned by the playability expert.

The inspectors used playability heuristic sets differently in this task compared to how they were using them during the evaluation sessions. The inspectors more frequently assigned multiple heuristics to describe the problem. In group H06, the inspectors assigned on average 1.21 heuristics, and in group H08, 1.13 heuristics on average were assigned for a problem. Another observation is that the division of the assigned heuristics increased, especially in group H08. In this task, the median value for how many heuristics were assigned to describe a problem was five. When you consider that the playability heuristic set that group H08 used included only 10 heuristics, it indicates that the inspectors had a hard time assigning the violated heuristic consistently.

When the violated heuristics that the inspectors assigned for playability problems were compared to heuristics that the playability expert assigned, the results are somewhat better than in the evaluation task. We calculated positive correspondence if any of the assigned heuristics matched with the heuristic assigned by the expert. In group H06, 48% of the inspectors had assigned the same violated heuristic as the playability expert on average. In group H08, the corresponding score was 41%. It is noticeable that the inspectors agreed quite consistently with the expert on the violated heuristic when looking at different game features. However, there is one exception, the Character Creation feature. Approximately half of the inspectors in group H06 had assigned same violated heuristic as the playability expert, but in group H08, only 28% of the inspectors had assigned the same heuristic. One probable reason for this is that the playability heuristic set that group H08 used is missing heuristics that would describe problems in the UI design (e.g. navigation issues). Although some inspectors agreed with the playability expert that a violated heuristic cannot be assigned, the majority of the inspectors had still assigned one.

Table 4 Summary of violated heuristic assignment during the problem analysis

		Group H06			Group H08		
Game Feature	Category (count)	Violated Heuristics	Assigned Heuristic	Agree with Expert	Violated Heuristics	Assigned Heuristics	Agree with Expert
Character Creation	User Interface (11) Gameplay (2)	4	1,16	52 %	6	1,10	28 %
Music Selection	User Interface (6) Gameplay (2) Mobility (1)	4	1,08	25 %	6	1,09	29 %
Dancing	User Interface (17) Gameplay (5)	5	1,29	55 %	4	1,16	54 %

4.3 Subjective Ratings of the Heuristic Sets

In addition to collecting hard evidence about identified playability problems in the game and how the inspectors assigned violated heuristics to describe the problems, we also wanted to know how the inspectors perceived the playability heuristic sets they used, and whether they found them useful during the game evaluation and problem analysis tasks. We were interested in knowing if there was any difference in opinion between the tasks and the groups. Twice during the experiment, the inspectors filled in a questionnaire in which we tried to find answers to these questions.

Sim et al. propose that, when the validity of the heuristics is evaluated, certain criteria such as correctness and coverage should be assessed [19]. Correctness means that the terminology that is used in the heuristic descriptions provides sufficient information for the inspectors. Coverage relates to the extent to which the heuristics represent the domain being evaluated. The correctness of the heuristics was measured by asking how the inspectors perceived the headings and the descriptions of the heuristics, and whether they were understandable. The coverage of the heuristics was measured to establish to what extent the inspectors were able to find a proper heuristic to describe the problem.

In both groups, the inspectors' opinion was that the headings of the playability heuristics helped them to assign a proper heuristic for a problem. A Wilcoxon Signed Rank test revealed no statistically significant difference in the usefulness of the headings, z=-.575, p=.565, with a small effect size (r=.13). The median score on the usefulness of the headings remained same in the evaluation task (Md=5) and the problem analysis task (Md=5).

When analyzing how easy the heuristic descriptions were to understand, the inspectors' opinion decreased slightly between the tasks, especially in group H06. A Wilcoxon Signed Rank Test revealed nearly a statistically significant reduction in group H06, z=-1.890, p=.059, with a large effect size (r=.051). The median score on the understandability of the descriptions decreased from the game evaluation (Md=6) to problem analysis (Md=5). However, the median score is still clearly on the positive side. In group H08, there was no statistically significant change between the tasks, z=-1,265, p=.206, with a small effect size (r=.27) even though the median score decreased as well from the game evaluation task (Md=6) to problem analysis task (Md=5).

The reason why the inspectors' opinion decreased between the tasks might be that the problem analysis task was more difficult for the inspectors, and using the heuristic sets also felt more difficult because they had to compare both the problem description and the heuristic description in order to determine the violated heuristic. In the game evaluation task, the inspectors had more information available to determine the violated heuristic. The inspectors knew the context in which the problem occurred, and they could determine whether the problem is related to the UI or gameplay. In the problem analysis task, the inspectors had to imagine a situation in which the problem appears, and they had to

rely on their memory and previous experience of the game, because the mobile phone or the games were not available.

The coverage of the heuristic set was measured by asking the inspectors how they perceived the heuristic sets in terms of finding a proper heuristic to describe a problem. When comparing questionnaire results after the game evaluation task, a Mann-Whitney U test revealed no significant difference in finding a proper heuristic for a playability problem in group H06 (Md=5, n=15) and group H08 (Md=4.5, n=18), U=113.5, z=-.793, p=.428, r=.13. Three inspectors from group H06 did not return the post-test questionnaire, and therefore n=15. The result is in line with an observation that the accuracy of assigning violated heuristics was almost equal. In addition, the inspectors did not know that only a third of the violated heuristics were the same when comparing them to heuristics assigned by the playability expert.

After the problem analysis task, a Mann-Whitney U test revealed a close to significant difference in finding a proper heuristic for a playability problem in group H06 (Md=4, n=7) and group H08 (Md=3, n=11), U=21, z=-1.615, p=.106, r=.38. The inspectors in group H06 more consistently assigned a violated heuristic when it was compared to heuristics that the playability expert assigned. The correspondence was 48% (during the evaluation task, it was 35%). In group H08, it was 41% and 33%, respectively.

Another observation from the problem analysis task is that the inspectors could not always find a proper heuristic to describe a problem. Especially the inspectors in the group H08 left the violated heuristic unassigned more often than the inspectors in group H06. The violated heuristic was left unassigned 23 times in group H06, whereas in group H08, it was done 89 times. The difference between the groups is significant, and this should presumably also be visible in the subjective assessment of the heuristic sets. However, the inspectors did not perceive the use of the playability heuristic sets as problematic in this respect. A Mann-Whitney U test revealed no significant difference in not finding a suitable heuristic for a playability problem in group H06 (Md=6, n=7) and group H08 (Md=7, n=11), U=30.5, z=-.781, p=.479, r=.18. Both groups agreed that for some playability problems, a violated heuristic could not be assigned, but the question of whether this disturbed the inspectors in the groups was not really answered.

The number of assigned heuristics to describe the playability problems increased in the problem assignment task, and the inspectors in both groups more often assigned multiple heuristics to describe the problem. During the evaluation task, the inspectors assigned 1.08 heuristics on average, whereas in the problem analysis, task group H06 assigned 1.21 heuristics and group H08 assigned 1.13 heuristics on average. A Wilcoxon Signed Rank Test revealed a statistically significant decrease in unambiguity of the heuristic descriptions in group H06 after using them in problem analysis task, z=-2,041, p=.041, with a large effect size (r=.55). The median score on the unambiguity of the heuristic descriptions decreased from the game evaluation task (Md=5) to the problem analysis task (Md=3). In group H08, there was no statistically significant change between the tasks.

The inspectors' opinion probably changed, because they had a hard time understanding the context of the playability problem properly. In one case, the playability problem was related to the lack of alternative game controls that could be used to control the game avatar. This is clearly a user interface problem, but some inspectors had interpreted this to mean that the game does not support different playing styles, which is related to the gameplay aspect of the game.

5. DISCUSSION

In this experiment, we have studied how novice inspectors use playability heuristic sets when they conduct a game evaluation and assign violated heuristics for identified playability problems. The assigned heuristics were compared to heuristics that a playability expert assigned for the same problems. We selected two playability heuristic sets that are quite different in their scope. Pinelle et al. have defined heuristics that mainly concentrate on usability issues [11], whereas Korhonen and Koivisto have defined heuristics that cover the user interface, gameplay, and mobility aspects of the game [10].

In the game evaluation task, the reporting activity of the groups was quite similar, and both groups produced several playability problem reports during the evaluation. In that sense, the groups were balanced, and there was no indication that inspectors in one group would have been more experienced or more active when conducting the evaluation. Assigning a proper heuristic to describe the playability problem depends quite a lot on how a problem is described, and what aspects are highlighted in the description. The effect of problem description was clearly visible in the experiment. In the game evaluation task, the inspectors identified the problems themselves and assigned a violated heuristic based on the description that they produced. Even though the problem reports described the same underlying problem, violated heuristics were in many cases very different and similarity to violated heuristics assigned by a playability expert was approximately 34%. In the problem analysis task, the playability problem descriptions were produced by a playability expert, and the inspectors assigned violated heuristics for them. The consistency of the violated heuristics increased to 48% in group H06, and in group H08, it was 41%. Our results support Somervell's and McCrickard's finding that by providing a set of usability problems to the inspectors, it is possible to more accurately determine the applicability of a heuristic set [15].

However, it should be noted that approximately half of the heuristics were assigned to describe a problem in group H08. In the worst case, eight out of ten heuristics were assigned as a violated heuristic, but none of them actually described the problem. This indicates that either the inspectors in group H08 did not understand the problem description, or the descriptions of the heuristics were not clear. In group H06, the inspectors also assigned more violated heuristics to describe problems, but the median value increased from 4 to 4.5 heuristics.

Although the accuracy of assigning a violated heuristic to playability problems improved in the problem analysis task, it should be noted that inspectors in group H06 perceived use of the heuristic set more difficult than in the evaluation task. Even though inspectors' opinions about the easiness remained clearly on the positive side, the reduction was remarkable. One probable reason for this is that the inspectors did not have enough material available that would have helped them to understand the playability problems. It was not possible to provide the inspectors with the mobile phone and the game, and they had to rely on their own memory and previous experience with the game when analyzing the problem descriptions. It was also possible that the inspectors did not encounter a situation in which the problem occurred, meaning that they had very little information available.

In the future, we should ensure that the inspectors have some background material available when analyzing problem descriptions. At a minimum, it could be a screenshot or a video showing the situation in which the problem occurs, but the best option would be that the inspectors would always get hands-on

experience of the game. This would help in analyzing the underlying structures of why the problem exists. Hvannberg et al. had also found that usability problems should be linked to the context for richer problem description [20].

We could of course ask if it really matters which heuristic is assigned to describe a playability problem, as long as the inspectors are identifying the problems that can have a negative influence on the gaming experience and should be fixed before the game is launched. In the professional game development projects it does, because it is easier for the developers to understand why an issue is a problem when they can refer to a correct playability heuristic which will describe the problem on a general level. In addition, developing playability heuristics that are easy to use and understandable will help inspectors to conduct the evaluation and make the evaluation session more efficient.

6. CONCLUSIONS

Research of playability heuristics to support usability inspection methods in videogame evaluations has been active in recent years, and there are multiple playability heuristic sets available. However, they differ quite a lot from each other and it is still unknown how well they support evaluation and help to describe identified playability problems in a game. We conducted an experiment in which 36 novice inspectors were divided into two groups, and they evaluated a mobile game using one of the two playability heuristic sets. In the second task, the inspectors analyzed playability problems that were collected from the game and assigned heuristics to describe the problems. The results show that a playability heuristic set needs to cover the main aspects of playability in order to be effectively used during the evaluation. The results of this work will help in developing playability heuristics that make expert review an effective method for evaluating videogames.

7. ACKNOWLEDGMENTS

We thank Dean Bent and Antti Lyyra for their generous help in arranging the devices and the game for the study, and Jenni Anttonen for her help in recruiting inspectors.

8. REFERENCES

[1] Korhonen, H., Paavilainen, J. and Saarenpää, H. 2009. Expert Review Method in Game Evaluations – Comparison of Two Playability Heuristic Sets. In proceedings of the *Academic MindTrek 2009 Conference*. ACM Press. 74-81. DOI= http://doi.acm.org/10.1145/1621841.1621856

[2] Fullerton, T., Swain, C. and Hoffman, S. 2004. *Game Design Workshop: Designing, Prototyping, and Playtesting Games*. CMP Books

[3] Rouse, R. 2001. *Game Design: Theory and Practice*. Wordware Publishing

[4] Schell, J. 2008. *The Art of Game Design*. Morgan Kaufmann

[5] Nielsen, J. and Molich, R. 1990. Heuristic evaluation of user interfaces. In proceedings of the *SIGCHI Conference on Human Factors in Computing Systems*. ACM. 249-256. DOI= http://doi.acm.org/10.1145/97243.97281

[6] Johnson, D. and Wiles, J. 2003. Effective affective user interface design in games. *Ergonomics* 46, 13-14, 1332-1345. DOI= http://dx.doi.org/10.1080/00140130310001610865

[7] Federoff, M.A. 2002. *Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games*. Master Thesis. Department of Telecommunications. Indiana University

[8] Nielsen, J. 1994. *Heuristic Evaluation*. In Usability Inspection Methods. Wiley & Sons, 25-62

[9] Desurvire, H., Caplan, M. and Toth, J.A. 2004. Using heuristics to evaluate the playability of games. In proceedings of the *SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM Press. 1509 - 1512. DOI= http://doi.acm.org/10.1145/985921.986102

[10] Korhonen, H. and Koivisto, E.M.I. 2006. Playability Heuristics for Mobile Game. In proceedings of the *International conference on Human computer interaction with mobile devices and services*. ACM Press. 9-16. DOI= http://doi.acm.org/10.1145/1152215.1152218

[11] Pinelle, D., Wong, N. and Stach, T. 2008. Heuristic evaluation for games: usability principles for video game design. In proceedings of the *SIGCHI Conference on Human Factors in Computing Systems*. ACM Press. 1453-1462. DOI= http://doi.acm.org/10.1145/1357054.1357282

[12] Jegers, K. 2008. Investigating the Applicability of Usability and Playability Heuristics for Evaluation of Pervasive Games. In proceedings of the *International Conference on Internet and Web Applications and Services*. 656-661. DOI= http://dx.doi.org/10.1109/ICIW.2008.54

[13] Nielsen, J. 1994. Enhancing the explanatory power of usability heuristics. In proceedings of the *SIGCHI Conference on Human Factors in Computing Systems*. ACM. 152-158. DOI= http://doi.acm.org/10.1145/191666.191729

[14] Mankoff, J., Dey, A.K., Hsieh, G., Kientz, J., Lederer, S. and Ames, M. 2003. Heuristic evaluation of ambient displays. In proceedings of the *SIGCHI Conference on Human Factors in Computing Systems*. ACM. 169-176. DOI= http://doi.acm.org/10.1145/642611.642642

[15] Somervell, J.P. and McCrickard, D.S. 2004. Comparing Generic vs. Specific Heuristics: Illustrating a New UEM Comparison technique. In proceedings of the *Human Factors and Ergonomics Society*. 2480 - 2484.

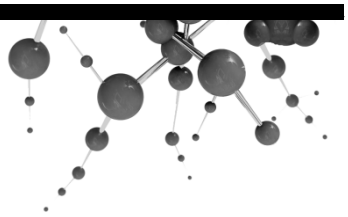
[16] Pagulayan, R.J., Keeker, K., Wixon, D., Romero, R.L. and Fuller, T. 2003. *User-centered Design in Games*. In Handbook for Human-Computer Interaction in Interactive Systems. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 883-906

[17] Hartson, H.R., Andre, T.S. and Williges, R.C. 2001. Criteria For Evaluating Usability Evaluation Methods. *International Journal of Human-Computer Interaction* 13, 4, 373 - 410.

[18] Korhonen, H. 2010. Comparison of playtesting and expert review methods in mobile game evaluation. In proceedings of the *3rd International Conference on Fun and Games*. ACM. 18-27. DOI= http://doi.acm.org/10.1145/1823818.1823820

[19] Sim, G., Read, J. and Cockton, G. 2009. Evidence Based Design of Heuristics for Computer Assisted Assessment. In proceedings of the *IFIP TC 13 International Conference on Human-Computer Interaction*. Springer. 204-216. DOI= http://dx.doi.org/10.1007/978-3-642-03655-2_25

[20] Hvannberg, E.T., Law, E.L.-C. and Lárusdóttir, M.K. 2007. Heuristic evaluation: Comparing ways of finding and reporting usability problems. *Interacting with Computers* 19, 2, 225-240. DOI= http://dx.doi.org/10.1016/j.intcom.2006.10.001

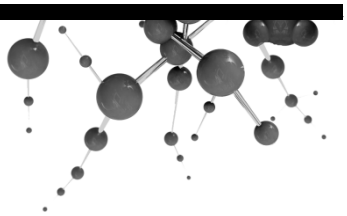


Paper VII

Korhonen, H. (forthcoming). Playability Heuristic Set Influences Inspectors' Performance in a Mobile Game Evaluation,

In review for publication in *Entertainment Computing Journal*

This article is not yet available. Pages 257 – 280 are not shown.



Appendix A:

Code	Game Usability Heuristics
GU1a	**Audio-visual representation supports the game**
	Game graphics should support gameplay and story and be informative for the player. In addition, the graphical look and feel should be consistent throughout the game. Audio can be used to evoke emotions and increase immersion. A good sound environment in the game supports a positive gaming experience. The graphics or audio should not prevent the player from performing actions or make them unnecessarily difficult.

GU1b	**A view to the gameworld supports smooth interaction and the camera behaves correctly**
	The view to the gameworld defines how the player perceives that gameworld and how well the player can immerse in the game events. The gameworld can be presented either in a 2D or a 3D view and there is no single best option of perspective, but it depends on the type of game and the game concept. It is more important to notice how the perspective influences the player's ability to interact in the gameworld.

GU2	**Screen layout is efficient and visually pleasing**
	The layout should present all necessary information to the player, but on the other hand, if the screen is filled with all kinds of information, it starts to look crowded. It is important that the player finds the navigation controls and they should not be mixed with the information that needs to be visible on the screen.

GU3	**Device UI and game UI are used for their own purposes**
	It should always be noticeable whether the player is dealing with the game user interface or device functions. The game interface should not use the device's user interface widgets in the game interface, because it breaks the immersion. The most impressive immersion is achieved when the game uses full-screen mode hiding other features.

GU4	Indicators are visible
	The player should see the information such as the current state of the game and the status of the avatar that is required for being able to play the game. Information that is frequently needed should be visible to the player all the time — if possible.

GU5	The player understands the terminology
	The terminology that is used in the game should be understandable and not misleading or unfamiliar to the players. Technical jargon should be avoided. For instance, terminology that is related to the game concept or features that the game needs from the device should be translated into more understandable language.

GU6	Navigation is consistent, logical, and minimalist
	Navigation consists of the game menu and the gameworld. The game menu consists of settings and selections for the desired game session. Different functions should be organized reasonably and possibly on different screens. However, long navigation paths should be avoided. Short navigation paths provide more clarity and are easier to remember. In the main game menu, the player should be able to start a game and have access to other important game features. In the gameworld, navigation should be intuitive and natural. Regardless of the complexity of the gameworld, players should be able to navigate there smoothly. With a proper set of control keys, navigation can be very intuitive and almost invisible.

GU7	Game controllers are consistent and follow standard conventions
	Using common conventions in game controllers reduces the time that is needed to learn to play the game since the player can use his or her knowledge from other games. Game controllers usually have specific keys for certain actions and every game should follow them. The game should also provide alternative game controllers because the players should be able to select their preferable controllers.

GU8	Game controls are convenient and flexible
	Novice players usually need only a subset of the controls when they start playing the game. On the other hand, veteran players often appreciate shortcuts and more advanced commands. It should be possible to customize the game controls or use shortcuts or macros. However, using shortcuts should not provide a major edge in a competitive player vs. player game. The configurability and amount of controls needed to play the game should be kept at a minimum, but they need to be sufficient. In addition, the controls should be designed according to the device's capabilities.

GU9	The game gives feedback on the player's actions
	A good user interface has a low response time on the player's actions. An action can be either a single key press or a more complicated input sequence. The player should notice immediately that the game has recognized the action by providing feedback. The most common way of providing feedback is to present it graphically. Other alternatives are to use audio or tactile feedback. Providing only auditory feedback is not acceptable since a player may be playing the game without sounds. Although the game needs to respond immediately to the player's actions, the consequences of the action can be shown to the player later. If an action cannot be performed immediately, the game should notify the player about the delay. Although the game needs to respond immediately to the player's actions, the consequences of the action can be shown to the player later. If an action cannot be performed immediately, the game should notify the player about the delay.

GU10	The player cannot make irreversible errors
	The game UI should confirm actions that can cause serious and irreversible damage, which affects the player's ability to play the game. Such errors are typically done in a shell menu or in the setting dialogs to the game character or the player's progress in the game. When mistakes happen, it is helpful to enable recovery.

GU11	The player does not have to memorize things unnecessarily
	The game should not stress the user's memory unnecessarily, unless it is part of the gameplay.

GU12	The game contains help
	Players do not often read manuals. Instead, the game should teach the player what he or she needs to know to start playing the game. This can be done through a tutorial mode at the beginning of the game. The tutorial mode should be divided into chapters that teach a couple of things at the beginning. Ideally, the tutorial could be embedded completely in the game so that help would be provided every time when it is really needed. Help is also often needed in error situations. If the game provides useful error messages, the player can understand better what caused the problem.

Code	Gameplay Heuristics
GP1	**The game provides clear goals or supports player-created goals**
	The player needs to understand the goals that exist in the game. The goals can be either set by the game or created by the players. The game should contain both short-term and long-term goals. Short-term goals provide repeated opportunities for reinforcement and keep players motivated to play the game. Long-term goals are usually more difficult to achieve and they can consist of several short term goals.

GP2	**The player sees the progress in the game and can compare the results**
	The players should have enough information so that they can see their progress towards the goals in the game. The progress can be shown to the player explicitly or implicitly. The players feel more motivated if they can compare themselves to other players or their previous achievements.

GP3	**The players are rewarded and the rewards are meaningful**
	The players should receive a meaningful reward as they progress in the game. In addition, the reward should be adjusted to the challenge that the player had to face in order to get it. The rewards schedule should be varying and frequent, but still unpredictable.

GP4	**The player is in control**
	The players want to be in control of what is happening in the gameworld. The players should be able to decide on actions they want to take and these actions should have an influence on the gameworld. If full control is not possible, the game should provide at least an illusion of control to the player.

GP5	**Challenge, strategy, and pace are in balance**
	The game should be designed so that the challenge is comparable to the players' current skills, then the players do not feel frustrated or bored with the game. In single-player games, the player can often choose the difficulty level and thus affect the challenge. The players learn new strategies as they play the game. There should not be dominating strategies for any part of the game. The pace should be adjusted to the game style and it can be intensive or deliberate. The game should allow the player to take a deep breath once in a while.

GP6	**The first-time experience is encouraging**
	The first impression of the game is formed within a few minutes and it is very difficult to change. The players should feel that they have learned the basics and have accomplished something. The first play session should make the player desire the next play session.

GP7	**The game story, if any, supports the gameplay and is meaningful**
	Even though the story plays an important role in many games, it should not dominate the gameplay. Some games do not even have or need a game story. If the game has a story, it should fit the other elements in the game and sound plausible to the player. Dialogue with non-player characters (NPC) should be meaningful and interesting to the player.

GP8	There are no repetitive or boring tasks
	A game should not repeat tasks without changing any conditions. Often, this repetition happens when the player needs to reach a certain goal before the game becomes interesting or challenging. During the training phase (tutorials), it is useful to repeat certain tasks so that the player learns and practices, for example, how the character is controlled in the game.

GP9	The players can express themselves
	The players should be able express themselves by, for instance, customizing their characters, acting in a certain way, or modifying the gameworld. Allowing the players to customize and personalize their game characters makes it more probable that they feel attachment to a game.

GP10	The game supports different playing styles
	The players vary a lot in terms of experience and preferred playing style. One of the main differentiating factors is the players' attitude towards risk-taking. The story can also influence whether the player wants to play as a hero or a villain. The player types will also determine how the players prefer to interact with the gameworld and with other players. In very simple games, different play styles are usually not supported and all players will have only one role.

GP11	The game does not stagnate
	The players should always feel that it is possible to reach the goals and the game progresses towards the goals. Game items or the balance of power should not counterbalance each other to result in an infinite loop of meaningless actions. The game should recognize immediately when the game is over and inform the players. The ending of the play session should be clearly indicated and restarting the game should be possible.

GP12	The game is consistent
	The gameworld and actions should be consistent and logical to the player. If something works in the beginning, the player assumes that it also works later on. Correspondingly, if the player is able to perform a certain action in the gameworld or for a game item, the player assumes that a similar kind of action is possible for other similar objects or in a similar situation as well. More actions can become available as the player progresses through the game.

GP13	The game uses orthogonal unit differentiation
	Each game item should have a purpose in the gameworld and it should be notably different from other similar game items and preferably restricted in some way. If the player needs to select character classes or roles in the game, they should be functionally different. In addition, interaction between the avatar and the non-player characters should support different interaction styles.

GP14	The player does not lose any hard-won possessions
	The game should maintain the possessions that the player has earned while playing the game and the player cannot lose them accidentally. However, in some cases the game can provide very high risks and the player can stake valuable game items which can be lost during the gameplay.

Code	Mobility Heuristics
MO1	**The play sessions can be started quickly**
	Starting a play session should be quick and easy because it is often used to fill up micro-breaks. The player should be able to skip introduction sections or startup screens. The game menu should be designed so that frequently used actions are not hidden behind a long navigation path. The default settings for the control keys and UI customization should be feasible for most of the players. In addition, any changes to the settings should be stored. The player should be able to continue play sessions from the point where the player stopped during the previous play session.

MO2	**The game accommodates the surroundings**
	Playing the game in mobile contexts should not disturb non-players in the vicinity. A game should provide means for conveniently adjusting the volume level or muting the game. Alternatively, the game should respect the profile settings of the device or ask whether the player wants to disable the audio features of the game.

MO3	**Interruptions are handled reasonably**
	There can be both internal and external interruptions, which will disturb the gameplay. Internal interruptions include incoming calls and received messages. External interruptions are related to the player's task switching in the current context or other unexpected events, which cannot be anticipated. If an interruption happens, the game should pause automatically and allow the player to continue gaming later on.

MO4	**The graphical design is accommodated to current brightness (Supplements GU1a)**
	The game settings should provide means to compensate changes in the environmental context. The players should be able to adjust screen brightness and color scheme in the game if necessary.

MO5	**The player should be aware of some device features while playing (Supplements GU3 and GU4)**
	The player should be aware of certain device functions when playing games with a mobile device. The most important information is network connection, battery, and modality of the keypad. Showing a clock will help the player to estimate when it is time to end the play session. Information should be presented using the same graphical style and user interface widgets that are used in the game user interface.

MO6	**Mobile devices have their own conventions for input (Supplements GU7)**
	Game design should follow the input conventions of the target device. For mobile devices with a physical keypad the keys have specific functions like a selection key. For touch screen devices, the conventions are related to gestures.

MO7	The tutorial should respond to immediate demand (Supplements GU12)
	For mobile games, the tutorial should be tailored to present only information that a player needs during the current play session. The tutorial could be divided into several subsections which will span across several play sessions.

Code	Multiplayer Heuristics
MP1	**The game supports communication**
	Communication is one of the cornerstones in multi-player games and, depending on the game, communication channels for different purposes need to be supported. In-game communication can be either synchronous or asynchronous. Chatting is the most frequently used method for in-game communication, but it can be restricted to cover certain areas in the gameworld or certain players. Asynchronous messaging is used with players who are not currently online.

MP2	**There are reasons to communicate**
	Communication is an essential part of social interaction. When players are either collaborating or competing this will generate discussions. Moreover, if players see how other players are doing in the game it will also generate discussion topics. The communication topics usually range from game related topics to common conversation to kill time during boring play periods.

MP3	**The game supports groups and communities**
	The game should support both short-term groups and long-lasting communities because they will increase social interaction and keep players playing the game. Short-term groups need to support communication and care taking of other members of the group. Long-lasting communities are more persistent groups which help players to advance in the game and increase social interaction.

MP4	**The game helps the player to find other players and game instances**
	Teaming up with other players is an essential aspect in multi-player games. The players should have a sense of the presence of other players and be able to find them. The game can direct players with similar levels to the same area in the gameworld where the players will meet each other. Alternatively, the game can automate matchmaking and help players find other players or game instances.

MP5	**The game provides information about other players**
	Knowing other players' online status, level and rank will help in planning and coordinating actions in the game. The information can be used when deciding whether to interact with some player and if the player can be trusted. Being aware of other players in the gameworld will increase the social aspects of the game.

MP6	**The design overcomes the lack of players and enables soloing**
	There are always situations when other players are not available. There may not be enough players in the gameworld or a player might play the game at times when there are not many other players around. There might also be some personal reasons to play alone. The player might have insufficient language skills to communicate with other players or they might want to try to survive by themselves in the game, especially on the lower levels. Whatever the reason is, the multiplayer game should also provide content for solo players.

MP7	The design minimizes deviant behavior
	The game design should minimize deviant behavior of the players, which is a common problem in multi-player games. Some players try to purposefully disrupt other players' gaming experience with their own behavior. Players can cheat, exploit or hack the game system or they harass other players intentionally. Game designers should pay special attention to minimizing deviant behavior and implement mechanisms that prevent such actions.

MP8	The design hides the effects of the network
	In online games, the effect of the network may become an issue for positive player experience. Updates to the gameworld should always happen without delays, but latency can disrupt the gameplay and cause jitter in real-time interaction. As the players can move in the physical environment while playing mobile games, they can be unintentionally disconnected from the game. The game design should minimize disturbances to the player or other players in the game.

MP9	Players should play with comparable players (Supplements GP6)
	In multi-player games the player population can be diverse. Some players have played longer than others and are thus more experienced in the gameworld and the tactics. They also have more developed avatars. The game design should help novice players to get familiar with the game and let them practice their skills and develop their avatar apart from the more experienced players during the first play sessions.

Code	Context-Aware Heuristics
CA1	**Perception of the current context**
	Discovering the correct context and playing the game when the context is favorable for a player is one source of fun in pervasive mobile games. The challenge, however, is that the game system and the player may not have a mutual understanding of the current context. In addition, the player may be uncertain of how the game system interprets the current context even though it was clearly observable. The game system should only use information that is collected from the player's current location and avoid rigid thresholds. The game system can also notify the player of what the current context is, if it does not contaminate the game experience.

CA2	**Players should have an equal chance to play**
	Utilization of context information in the game system should be carefully designed and all players should have an equal possibility to access relevant information. The dynamic nature of the context or player-related reasons might cause some context information to become unreachable. The game should be designed so that the player progression is not dependent on context information that is unreachable by many players.

CA3	**Adjustable play sessions**
	Play sessions of pervasive mobile games are often blended into other activities of the player. This will cause the situations where the play sessions are fragmented and the player might have to quit playing the game without prior notice. The game should be designed so that the chance of winning the game is not dependent on the player's ability to play the game. In addition, the player might need to adjust the pace of the game and match it to available playing time. Sometimes it is preferable to play the game almost in real time while some other times a slower pace is preferred.

CA4	**Communication outside the gameworld (Supplements MP1)**
	Pervasive mobile games expand the need for communication outside the gameworld. The players need to be aware of game events even though they are offline. Offline communication should not be overwhelming, because it can lead to negligence and disturb the player's other activities. The game events that require the players' actions should be delivered immediately, while others can be postponed and delivered as a summary type of messages.

Publications in the **Dissertations in Interactive Technology** series

Details of the dissertations are available at
http://www.uta.fi/sis/tauchi/dissertations.html.

1. **Timo Partala:** Affective Information in Human-Computer Interaction
2. **Mika Käki:** Enhancing Web Search Result Access with Automatic Categorization
3. **Anne Aula:** Studying User Strategies and Characteristics for Developing Web Search Interfaces
4. **Aulikki Hyrskykari:** Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid
5. **Johanna Höysniemi:** Design and Evaluation of Physically Interactive Games
6. **Jaakko Hakulinen:** Software Tutoring in Speech User Interfaces
7. **Harri Siirtola:** Interactive Visualization of Multidimensional Data
8. **Erno Mäkinen:** Face Analysis Techniques for Human–Computer Interaction
9. **Oleg Špakov:** iComponent – Device-Independent Platform for Analyzing Eye Movement Data and Developing Eye-Based Applications
10. **Yulia Gizatdinova:** Automatic Detection of Face and Facial Features from Images of Neutral and Expressive Faces:
11. **Päivi Majaranta:** Text Entry by Eye Gaze
12. **Ying Liu**: Chinese Text Entry with Mobile Devices
13. **Toni Vanhala**: Towards Computer-Assisted Regulation of Emotions
14. **Tomi Heimonen:** Design and Evaluation of User Interfaces for Mobile Web Search
15. **Mirja Ilves:** Human Responses to Machine-Generated Speech with Emotional Content
16. **Outi Tuisku:** Face Interface
17. **Juha Leino:** User Factors in Recommender Systems: Case Studies in e-Commerce, News Recommending, and e-Learning
18. **Joel S. Mtebe**: Acceptance and Use of eLearning Solutions in Higher Education in East Africa
19. **Jussi Rantala:** Spatial Touch in Presenting Information with Mobile Devices
20. **Katri Salminen:** Emotional Responses to Friction-based, Vibrotractile, and Thermal Stimuli
21. **Selina Sharmin:** Eye Movements in Reading of Dynamic On-Screen Text in Various Presentation Formats and Contexts
22. **Tuuli Keskinen:** Evaluating the User Experience of Interactive Systems in Challenging Circumstances
23. **Ogunbase Adewunmi Obafemi:** Pedagogical Design and Pedagogocal Usability of Web-Based Learning Environments: Comparative Cultural Implications from Africa and Europe