**T.C.**

**GALATASARAY UNIVERSITY**
**GRADUATE SCHOOL OF SOCIAL SCIENCES**
**DEPARTMENT OF RADIO, TELEVISION AND CINEMA**

# ANALYZING PLAYER EXPERIENCE: MULTI-MODAL EVALUATION OF PLAYER EXPERIENCE IN DIFFERENT GAMING PLATFORMS

**DOCTORAL DISSERTATION**

**Çakır AKER**

**Supervisor: Prof. Dr. Kerem RIZVANOĞLU**

**JUNE 2018**

## ACKNOWLEDGEMENTS

Since I became an official PhD student at Galatasaray University's Graduate School of Social Sciences and its Media and Communication Studies Doctoral Program, I have worked with a great number of people who have offered me their support that can never be repaid. No words can express my profound gratitude I have for my supervisor, Prof. Dr. Kerem RIZVANOĞLU for his patience, guidance and faith in my ability to complete this journey. I would also like to thank him for his time and dedication towards my thesis and for paving a way for my academic and intellectual progression, in addition to this PhD dissertation. His vast knowledge on the fields of human-computer interaction, user experience and user research have expanded my horizon. My sincere appreciation also goes for Prof. Dr. Serhat GÜNEY and Asst. Prof. Dr. Güven ÇATAK for their supervision, support and invaluable guidance throughout the PhD period. I would also like to thank Asst. Prof. Dr. Yavuz İNAL for his support, guidance and contribution during the study. Moreover, I thank the gamers in Bahçeşehir University who participated in this research. Their willingness to share their time and insights made this study possible.

Finally, I thank the people who have a special place in my heart. I would like to thank my colleagues and friends for their constant support, help and friendship. Without their support, this dissertation would never have been finished. Lastly, but most importantly, I would like to thank my mother who has always been supportive and caring during the difficult times that I have been through. Her unconditional love and unfailing encouragement made me who I am today.

**Çakır AKER**

**June 2018**

# TABLE OF CONTENTS

# ABBREVIATIONS

**3D –** Three Dimensional

**CRT –** Cathode Ray Tube

**FPS –** First Person Shooter

**GEQ –** Game Experience Questionnaire

**HCI –** Human Computer Interaction

**MMORPG –** Massively Multiplayer Online Role Playing Game

**NPS –** Net Promoter Score

**PC –** Personal Computer

**PM –** Playability Model

**PX –** Player Experience

**QA –** Quality Assurance

**SUS –** System Usability Scale

**UX –** User Experience

**VR –** Virtual Reality

# LIST OF TABLES

## LIST OF FIGURES

## RESUME

Les jeux vidéo sont maintenant considérés comme l'un des principaux indicateurs des progrès technologiques. C'est un fait que, il y a plusieurs décennies, les jeux vidéo étaient limités par ce que la technologie actuelle pouvait offrir, mais ils ont maintenant la capacité de piloter eux-mêmes l'innovation technologique. Néanmoins, la concurrence sur le terrain a fait émerger différentes plates-formes de jeu. Des consoles aux téléphones mobiles, des jeux sont développés pour plusieurs de ces plateformes à la fois. Ces nouvelles capacités des plates-formes de jeu et des méthodes d'interaction pour les jeux vidéo ont mis en évidence l'importance de la conception centrée sur l'utilisateur dans les jeux. Fondamentalement, ces différentes plates-formes de jeu offrent des expériences différentes. Par conséquent, il est tout aussi important de comprendre l'importance des différences de plate-forme autant que les différents jeux en termes de compréhension de l'expérience des joueurs. Néanmoins, le domaine de la recherche se concentre souvent sur les jeux informatiques et tente de fournir un guide pour l'évaluation des jeux et des expériences tout en négligeant la nécessité d'analyser les différences entre les plateformes en termes d'expérience. De plus, il existe plusieurs méthodes d'évaluation pour analyser l'expérience du joueur. Cependant, aucune de ces méthodes ne s'est avérée efficace et inclusive lorsqu'elle est utilisée individuellement. Plus important encore, il a été observé que la majorité de ces études sur l'évaluation de l'expérience des joueurs n'ont pas été validées par divers tests portant sur différentes plates-formes et jeux. Comprendre les différences d'expérience des joueurs entre les plateformes de jeu dans une perspective inclusive est crucial pour identifier comment les joueurs interagissent avec le système et guider les concepteurs de jeux pour développer et évaluer leurs jeux qui répondraient aux attentes des joueurs.

Le but et l'objectif du travail de recherche est double. Tout d'abord, cette étude se concentre sur l'expérience des joueurs et la jouabilité des jeux à la lumière des différences de plate-forme, visant à indiquer quelles méthodes peuvent identifier les différences d'expérience des joueurs entre les plateformes. Cette thèse rapporte une série d'études sur les analyses des expériences des joueurs en relation avec les plateformes de jeu utilisées dans le jeu. Les définitions de l'expérience du joueur; les différences de jouabilité entre

les plateformes de jeu; la convivialité et les évaluations de la jouabilité sur les jeux; et diverses approches méthodologiques pour analyser les expériences des joueurs constituent la portée de cette étude. Ainsi, l'étude est fondamentalement un travail multidisciplinaire combinant la recherche de HCI et la recherche de jeu qui vise à analyser et inspecter diverses méthodes d'évaluation de la jouabilité et de l'expérience du joueur pour analyser les différences de plate-forme.

Deuxièmement, cette étude vise à identifier et présenter les différences d'expérience des joueurs entre plates-formes inclusivement par une évaluation multimodale, et se concentre sur la présentation d'un set méthodologique pour évaluer les différences de plateforme en termes d'expériences des joueurs, qui a le potentiel d'améliorer encore le domaine de la recherche ainsi que l'industrie du jeu. Pour ce faire, l'utilisation de diverses méthodes d'évaluation dans une perspective multimodale est inspectée pour évaluer l'efficacité de ces méthodes. Une revue complète de la littérature a été menée pendant l'étude, inspectant diverses méthodes et techniques d'évaluation de la jouabilité pour évaluer l'expérience des joueurs dans les jeux. L'étude identifie divers problèmes de jouabilité à travers différentes méthodes d'évaluation qui incluent une analyse basée sur l'utilisabilité, une évaluation empirique via des enquêtes, une évaluation analytique via des heuristiques, et des évaluations via le test de jeu. Pour évaluer les différences de plate-forme de manière holistique, trois études expérimentales consécutives utilisant différentes approches méthodologiques ont été menées ; évaluation basée sur l'utilisabilité, évaluation empirique et évaluation heuristique. Tous les tests ont été planifiés selon une approche multimodale qui a utilisé des playtests en combinaison avec des méthodes d'évaluation spécifiques pour permettre une analyse inclusive de l'expérience des joueurs. Pendant les tests, deux jeux occasionnels (Fruit Ninja pour la première étude et Plants vs Zombies pour les deux dernières études) ont été examinés.

Dans la première phase expérimentale de l'étude, un cadre conceptuel a été proposé pour indiquer différents éléments de jouabilité qui traitent des éléments de jeu «à l'écran» et «hors écran». Une plate-forme de console avec un périphérique Kinect et une plate-forme mobile ont été évaluées en termes de capacités de contrôle et d'interaction. En conséquence, le modèle de jouabilité basé sur les attributs qui définissent la facilité

d'utilisation et les heuristiques de mobilité ont été utilisés pour l'analyse. De plus, pour analyser l'expérience du joueur d'un point de vue inclusif, des tests d'utilisateurs comprenant des entrevues et des évaluations d'eye-tracking ont été utilisés. L'étude a fondamentalement fourni une approche basée sur les attributs d'utilisabilité. Les résultats de cette étude expérimentale indiquent que les attributs «efficacité», «apprentissage» et «motivation» ont été observés pour identifier les problèmes de jouabilité concernant la différence entre les plateformes. Cependant, mettant en évidence l'importance critique d'une approche multimodale pour les évaluations, le cadre proposé n'a pas fourni suffisamment d'indicateurs concernant les expériences des joueurs entre les plateformes lorsqu'elles sont appliquées individuellement.

Dans la seconde étude expérimentale, une évaluation empirique visait à analyser les différences d'expérience entre les plateformes. Game Experience Questionnaire (GEQ) a été utilisé afin d'évaluer l'expérience du joueur entre PC et plates-formes mobiles. Pour proposer une approche multimodale, des méthodes de test de jeu comprenant des entretiens semi-structurés, des observations et des enregistrements audiovisuels ont été utilisées. De plus, des échelles d'utilisabilité ont été administrées pour la collecte d'informations concernant l'utilisabilité du jeu. Comme dans l'étude précédente, il a été constaté que l'application du questionnaire ne fournissait pas de résultats indiquant des différences significatives entre les plateformes lorsqu'elles étaient utilisées individuellement. Les résultats suggèrent que seul un certain nombre d'items de l'enquête étaient capables d'identifier les différences d'expérience de jeu entre les plateformes, mais que l'application des échelles d'utilisabilité était efficace en termes d'approche méthodologique. Trois éléments du GEQ, «flux», «compétence» et «défi» et l'utilisation de la version modifiée de l'échelle d'utilisabilité du système (SUS) ont été jugés utiles pour analyser les différences entre les expériences des joueurs. De plus, les feedbacks de joueurs recueillis à travers les playtests ont démontré leur influence dans la présentation d'une compréhension en profondeur et ont ainsi présenté une perspective d'analyse des différentes attitudes et sentiments des joueurs.

Dans la troisième étude, une évaluation heuristique utilisant l'heuristique de la jouabilité a été utilisée pour analyser les expériences des joueurs entre les plateformes. PC

et plates-formes mobiles ont été évalués au cours de l'étude. En retraçant le travail précédent, des échelles d'utilisabilité et des playtests comprenant des interviews semi-structurées, des observations et des enregistrements audiovisuels ont été utilisés pour analyser l'expérience du joueur dans une perspective holistique. Il a été constaté que l'heuristique de la jouabilité ne fournissait pas de différences significatives en ce qui concerne les différences d'expérience entre les plates-formes lorsqu'elles sont appliquées par elle-même. Cependant, une sélection de 15 heuristiques de l'heuristique de la jouabilité ont été jugées admissibles pour identifier des problèmes de jouabilité plus spécifiques qui indiquaient des différences d'expérience entre les plateformes. Les feedbacks des joueurs se sont avérés inestimables pour l'analyse des différences d'expérience qui ont permis la catégorisation des problèmes de jouabilité dans une nouvelle perspective. L'identification et le regroupement des dépendances de problèmes de jouabilité proposées et de leurs associations avec des éléments heuristiques de jouabilité suggérés se sont révélés utiles pour identifier différentes expériences de joueurs entre plates-formes ainsi que pour augmenter l'applicabilité de la méthode. De plus, les échelles d'utilisabilité ont fourni des résultats encourageants pour l'analyse de l'expérience des joueurs.

Les résultats de ces études ont indiqué que chaque méthode révélait un aspect différent en ce qui concerne l'expérience et la jouabilité du joueur. Par conséquent, la conception itérative des études a permis une évaluation complète de l'expérience du joueur. Il a été déterminé que l'utilisation d'une méthode d'évaluation particulière ne donnait aucun résultat définitif en termes de comparaison des expériences des joueurs entre les plateformes. En conséquence, une approche multimodale a été indiquée comme une nécessité pour mener une analyse inclusive sur les expériences des joueurs. De plus, les résultats de chaque étude présentaient des éléments et des caractéristiques spécifiques qui permettaient d'évaluer l'expérience du joueur entre différentes plateformes d'un point de vue holistique. Ainsi, la triangulation des méthodes a été utile pour identifier les raisons des différences d'expérience des joueurs entre les plateformes inclusivement et pour la vérification croisée des résultats. Il a été constaté que l'application de différentes approches méthodologiques combinées à des playtests était une méthode efficace pour analyser l'expérience du joueur en détail.

En conclusion, l'applicabilité de l'ensemble méthodologique proposé a été inspectée à travers les phases du processus de développement du jeu, fournissant un modèle inclusif pour l'analyse des différences de plate-forme. Les étapes de production d'un jeu occasionnel ont été associées aux méthodologies et aux procédures d'évaluation pour guider davantage les chercheurs et les développeurs. Ces résultats peuvent être utilisés pour développer des ensembles de méthodologies supplémentaires pour évaluer différents genres de jeu ou plates-formes pour améliorer davantage le domaine. En appliquant le cadre de méthodologies proposé et leurs sous-items, la différence d'expérience entre les plateformes peut être mieux comprise.

# ABSTRACT

Video games are now considered to be one of the leading indicators of technological advancements. It is a fact that several decades ago, video games were limited with what current technological advancements could offer but now they even have the capability of steering the technological innovation on their own. Nevertheless, the growing competition in the field has caused different gaming platforms to emerge. From consoles to mobile phones, games are being developed for several of these platforms at once. These novel capabilities of gaming platforms and interaction methods for video games have highlighted the importance of user-centered design in games. Fundamentally, these different gaming platforms offer different experiences. Therefore, it is as crucial to understand the importance of platform differences as much as different games in terms of understanding player experience. Nevertheless, the research field often focuses on computer games and tries to provide a guide for the evaluation of games and experiences yet neglects the need for analyzing differences between platforms in terms of experience. Moreover, there are several evaluation methods for analyzing player experience. However, none of those methods proved to be effective and inclusive when used individually. Most importantly, it was observed that majority of these studies on evaluating player experience were not validated by diverse tests addressing different platforms and games. Comprehending the differences in player experience between gaming platforms from an inclusive perspective is crucial for identifying how the players interact with the given system as well as guiding game designers to develop and evaluate their games that would meet the player expectations.

The aim and the focus of the research work is twofold. Firstly, this study focuses on player experience and playability of games in the spotlight of platform differences, aiming to indicate which methods can identify player experience differences between platforms. This thesis reports on a series of studies on analyses of player experiences and playability evaluations in relation to the gaming platforms used in game play. The definitions of player experience; playability differences between gaming platforms; usability and the

playability evaluations on games; and various methodological approaches for analyzing the player experiences constitute the scope of this study. Thus the study is fundamentally a multidisciplinary work combining HCI research and game research which aims to analyze and inspect various playability and player experience evaluation methods for analyzing platform differences.

Secondly, this study aims to identify and present the differences in player experience between platforms inclusively by a multimodal evaluation, and focuses on presenting a methodological set for evaluating platform differences in terms of playability and player experiences which has the potential for further improving the research field as well as the gaming industry. To achieve that, utilization of various evaluation methods from a multi-modal perspective is inspected for evaluating the effectiveness of these methods. A comprehensive literature review was conducted during the study, inspecting various playability evaluation methods and techniques for evaluating player experience in games. The study addresses and identifies various playability issues through various evaluation methods which include a usability-based analysis, an empirical evaluation via surveys, an analytical evaluation via heuristics, and evaluations via playtesting. To evaluate platform differences in a holistic manner, three consecutive experimental studies employing different methodological approaches were conducted; usability based evaluation, empirical evaluation and heuristic evaluation. All the tests were planned following a multi-modal approach which playtests were employed in combination with specific evaluation methods to allow an inclusive analysis of player experience. During the tests, two casual games (*Fruit Ninja* for the first study and *Plants vs Zombies* for the latter two studies) were examined.

In the first experimental phase of the study, a conceptual framework was proposed for indicating different playability elements that address “on-screen” and “off-screen” game elements. A console platform with Kinect peripheral and a mobile platform were evaluated in terms of their distinct control and interaction capabilities. Accordingly, the Playability Model which is based on the attributes that define usability, and Mobility heuristics were utilized for analysis. Additionally, to analyze the player experience from an inclusive perspective, user-tests including interviews and eye-tracking evaluations

were employed. The study fundamentally provided an approach based on usability attributes. The results of this experimental study indicated that the attributes of “effectiveness”, “learnability” and “motivation” were observed to identify playability issues regarding the difference between platforms. However, highlighting the critical importance of a multi-modal approach for evaluations, the proposed framework did not provide sufficient indicators regarding player experiences between platforms when applied individually.

In the second experimental study, an empirical evaluation was aimed for analyzing player experience differences between platforms. Game Experience Questionnaire (GEQ) was utilized in order to evaluate the player experience between PC and mobile platforms. For proposing a multi-modal approach, playtesting methods including semi-structured interviews, observations and audio-visual recordings were additionally employed. Moreover, usability scales were administered for gathering information regarding the usability of the game. Similar to the previous study, it was found that the application of the questionnaire did not provide results which indicate significant differences between platforms when employed individually. The results suggested that only a number of survey items were capable of identifying player experience differences between platforms yet the application of usability scales were effective in terms of methodological approach. Three items of the GEQ, “flow”, “competence”, and “challenge” and the employment of the modified version of the System Usability Scale (SUS) were found useful for analyzing player experience differences. Moreover, player feedbacks gathered through playtests demonstrated to be influential in presenting an in-depth comprehension and thus presenting a perspective for analyzing different attitudes and feelings of the players.

In the third study, a heuristic evaluation utilizing Playability Heuristics was employed for analyzing player experiences between platforms. PC and mobile platforms were evaluated during the study. Tracing the previous work, usability scales and playtests including semi-structured interviews, observations and audio-visual recordings were used for analyzing the player experience from a holistic perspective. It was found that Playability Heuristics did not provide significant differences regarding the differences in player experience between platforms when applied by itself. However, a selection of 15

heuristics from Playability Heuristics were found to be eligible for identifying more specific playability issues which indicated experience differences between platforms. The player feedbacks proved to be invaluable for analyzing the experience differences which allowed the categorization of playability problems from a novel perspective. The identification and grouping of the proposed playability problem dependencies and their associations with suggested playability heuristic items proved to be useful for identifying different player experiences between platforms as well as increasing the applicability of the method. Furthermore, the usability scales provided supportive results for analyzing player experience.

The results of these studies indicated that each method revealed a different aspect regarding the player experience and playability. Consequently, the iterative design of the studies allowed a comprehensive evaluation of the player experience. It was identified that employing a particular evaluation method individually yielded no definitive results in terms comparing player experiences between platforms. As a result, a multi-modal approach was indicated as a necessity for conducting inclusive analysis on player experiences. Furthermore, results from each study presented specific items and features which allowed evaluating the player experience between different platforms from a holistic point of view. Thus, the triangulation of methods was useful for identifying the reasons of the player experience differences between platforms inclusively and for cross verification of the findings. It was found that the application of different methodological approaches combined along with playtests were an efficient method for analyzing the player experience in detail.

To conclude, applicability of the proposed methodological set was inspected through phases of game development process, providing an inclusive model for analyzing platform differences. The production stages of a casual game were associated with the methodologies and evaluation procedures for further guiding researchers and developers alike. These findings can be used to develop additional sets of methodologies for evaluating different game genres or platforms for further improving the field. By applying the proposed framework of methodologies and their sub items, the difference in player experience between platforms can be understood in a better way.

# ÖZET

Video oyunları günümüzde teknolojik ilerlemenin ön önemli göstergelerinden olarak kabul görmektedir. Birkaç on yıl öncesine kadar video oyunlarının teknolojik limitlere tabi olması söz konusuyken, artık oyunların teknolojik ilerlemede etkili olduğu gerçeği söz konusudur. Bununla birlikte, alandaki artan rekabet sebebiyle yeni oyun platformlarının doğuşu da söz konusu olmuştur. Konsollardan mobil telefonlara, oyunlar birden fazla platform için geliştirilmeye başlanmıştır. Platformların sunduğu bu yenilikçi oyun imkânları ve etkileşim metotları, oyunlarda kullanıcı odaklı tasarımın öneminin altını çizmiştir. Temelde bu farklı platformlar farklı deneyimler sunmaktadır. Bu sebeple oyuncu deneyimini anlamak için platform farklılıklarının ve farklı oyunların sunduğu deneyim farklılıklarının anlaşılması elzemdir. Buna rağmen araştırmalar sıklıkla bilgisayar oyunlarına odaklanmakta, oyun ve deneyim değerlendirmesi amacına dair yol gösterici sunmaya çalışmakta, ancak oyuncu deneyimi temelinde platform farklılıklarına bakmamaktadır. Benzer olarak, oyuncu deneyiminin analizi adına yalnızca birkaç değerlendirme yöntemi mevcuttur. Fakat bu yöntemlerden hiç biri tek başına uygulandığında etkili ve kapsamlı olamamaktadır. En önemlisi, bu çalışmaların çoğunda oyuncu deneyimi ölçümlerinin geçerliliğinin doğrulanmadığı görülmektedir. Oyuncuların belirlenen bir sistem üzerinden nasıl etkileşim kurduklarını ve oyun geliştiricilerin oyuncu beklentilerini karşılamak adına oyunlarını nasıl değerlendireceklerini tanımlayabilmek için, platformlar arası oyuncu deneyimini kapsamlı bir açıdan anlamak kritik önem taşımaktadır.

Çalışmanın amacı ve odağı iki yönlüdür. Öncelikle çalışma, oyuncu deneyimleri ve oynanabilirlik kavramlarına, platform farkları özelinde odaklanmakta ve hangi yöntemlerin platformlar arası farklara işaret edebileceğini belirlemeyi amaçlamaktadır. Bu tez çalışmasında, bir seri oyuncu deneyimi analizi çalışmasının platform farklılıkları ile ilgili ilişkiler göz önüne alınarak raporlanması söz konusudur. Oyuncu deneyiminin tanımlanması, oyunlarda oynanabilirlik ve oynanabilirlik değerlendirmeleri ile oyuncu deneyimini analiz etmek için kullanılan çeşitli yöntem ve yaklaşımlar, bu çalışmanın konusunu oluşturmaktadır. Bu nedenle çalışma temel olarak, platform farklılıklarını analiz etmek için çeşitli oynanabilirlik değerlendirme yöntemlerini analiz etmeyi ve incelemeyi

amaçlayan, insan-bilgisayar etkileşimi alanı ile oyun araştırmaları alanlarını birleştiren çok disiplinli bir yapıya sahiptir.

İkinci olarak bu çalışma, multimodal bir değerlendirme ile platformlar arasındaki oyuncu deneyimlerindeki farklılıkları tanımlamayı ve sunmayı amaçlamakta, araştırma alanını ve oyun endüstrisini daha da geliştirebilecek potansiyele sahip platform farklılıkları açısından oyuncu deneyimini inceleyen bir yöntem seti sunmaya odaklanmaktadır. Bunun için çalışmada, multimodal bir bakış açısıyla çeşitli değerlendirme yöntemlerinin kullanılması ve bu yöntemlerin etkililiğinin incelenmesi söz konusu olmuştur. Çalışma sırasında, oyunda oyuncu deneyimini değerlendirmek için çeşitli oynanabilirlik değerlendirme yöntemleri ve teknikleri incelenerek kapsamlı bir literatür taraması yapılmıştır. Çalışma, kullanılabilirlik temelli bir analiz, anket yoluyla ampirik bir değerlendirme, sezgisel yöntemle analitik değerlendirme ve oyun testi yoluyla yapılan değerlendirmeleri içeren çeşitli değerlendirme yöntemleri ile çeşitli oynanabilirlik konularını ele almakta ve tanımlamaktadır. Platform farklılıklarını bütünsel bir şekilde değerlendirmek için üç ardışık deneysel çalışma yapılmıştır; Kullanılabilirliğe dayalı değerlendirme, ampirik değerlendirme ve sezgisel değerlendirme. Tüm testler, oyuncu deneyimlerinin kapsamlı analizine olanak verecek şekilde, spesifik değerlendirme yöntemleri ve oyun testlerini kapsayacak biçimde planlanmıştır. Testler esnasında iki '*casual*' türde oyun (ilk çalışma için *Fruit Ninja* ve diğer iki çalışma için *Plants vs Zombies* oyunları) incelenmiştir.

Çalışmanın ilk deney aşamasında, "on-screen" ve "off-screen" oyun öğelerini ele alan farklı oynanabilirlik öğelerini göstermek için kavramsal bir çerçeve önerilmiştir. '*Kinect*' çevresel birimine sahip konsol platformu ile mobil platform, farklı kontrol ve etkileşim yetenekleri açısından değerlendirilmiştir. Buna göre, kullanılabilirliği tanımlayan niteliklere dayanan 'Playability Model' ve 'Mobility' sezgiselleri analiz için kullanılmıştır. Ek olarak, oyuncu deneyimini kapsayıcı bir bakış açısıyla analiz etmek için, görüşmeler ve göz izleme değerlendirmeleri de dahil olmak üzere kullanıcı testleri uygulanmıştır. Çalışma temel olarak kullanılabilirlik özelliklerine dayanan bir yaklaşım sağlamıştır. Bu deneysel çalışmanın sonuçları, "effectiveness", "learnability" ve "motivation" özelliklerinin, platformlar arasındaki oynanabilirlik farklarını belirlediğini

göstermiştir. Bununla birlikte, değerlendirmeler için uygulanan yöntem, bireysel olarak uygulandığında platformlar arasındaki oyuncu deneyimlerine ilişkin yeterli göstergeler sağlamadığını göstermiştir ve çok yönlü araştırmanın önemini vurgulamıştır.

İkinci deneysel çalışmada, platformlar arasındaki oyuncu deneyim farklılıklarını analiz etmek için ampirik bir değerlendirme amaçlanmıştır. PC ve mobil platformlar arasındaki oyuncu deneyimini değerlendirmek için 'Game Experience Questionnaire' (GEQ) kullanılmıştır. Multimodal yaklaşım önerebilmek için yarı yapılandırılmış mülakatlar, gözlemler ve görsel-işitsel kayıtlar da dâhil olmak üzere oyun testi yöntemleri de kullanılmıştır. Ayrıca, oyunun kullanılabilirliği ile ilgili bilgi toplamak için kullanılabilirlik ölçekleri uygulanmıştır. Bir önceki çalışmaya benzer şekilde, yöntemde belirtilen anketin, bireysel olarak kullanıldığında platformlar arasında önemli farklılıklar gösteren sonuçlar sağlamadığı bulunmuştur. Sonuçlar, sadece birkaç anket başlığının platformlar arasındaki oyuncu deneyim farklılıklarını tanımlayabildiğini, bununla birlikte kullanılabilirlik ölçeklerinin metodolojik yaklaşım açısından etkili olduğunu göstermiştir. GEQ başlıklarından "flow", "competence" ve "challenge" maddeleri ve Sistem Kullanılabilirlik Ölçeği'nin (SUS) çalışmada kullanılması, oyuncu deneyim farklılıklarını analiz etmek için faydalı bulunmuştur. Ayrıca, oyun testleri yoluyla toplanan oyuncu geri bildirimlerinin, derinlemesine bir analiz sunmada etkili olduğu ve bu nedenle oyuncuların farklı tutum ve hislerini analiz etmek için farklı bir bakış açısı sunduğu belirlenmiştir.

Üçüncü çalışmada, platformlar arasındaki oyuncu deneyimlerini analiz etmek için yapılan sezgisel değerlendirmede 'Playability Heuristics' sezgisel seti kullanılmıştır. Çalışmada PC ve mobil platform değerlendirilmiştir. Bir önceki çalışmaya benzer olarak, yarı yapılandırılmış görüşmeler, gözlemler ve görsel-işitsel kayıtlar da dâhil olmak üzere kullanılabilirlik ölçekleri ve oyun testleri oyuncu deneyimini bütünsel bir bakış açısıyla analiz etmek için kullanılmıştır. 'Playability Heuristics' sezgisel setinin, kendi başına uygulandığında, platformlar arasındaki oyuncu deneyimindeki farklılıkları belirleme konusunda yetersiz kaldığı gözlemlenmiştir. Bununla birlikte, sezgisel setinin 15 sezgiselinin platformlar arası deneyim farklılıklarını gösteren spesifik oynanabilirlik sorunlarını göstermede etkili olabildiği bulunmuştur. Oyuncu geribildirimlerinin, oynanabilirlik sorunlarının değerlendirilmesi ve bu sorunların yeni bir bakış açısıyla

kategorize edilebilmesi açısından çok değerli olduğu bulunmuştur. Belirtilen oynanabilirlik sorunlarının dayanakları ile sezgisel seti maddeleri arasında kurulan bağlar sayesinde oluşturulan bu kategorizasyon, oyuncu deneyimi açısından platform farklılıklarının ortaya konması ve uygulanabilir bir yaklaşım sunması açısından önerilmiştir. Ayrıca, kullanılabilirlik ölçekleri, oyuncu deneyimini analiz etmek için destekleyici sonuçlar sağlamıştır.

Bu çalışmaların sonuçları, her bir yöntemin oyuncu deneyimi ve oynanabilirliği ile ilgili farklı bir yönünü ortaya koyduğunu göstermiştir. Sonuç olarak, çalışmalarda ortaya konulan iteratif ve birbirini takip eden araştırma yaklaşımı, oyuncu deneyiminin kapsamlı bir değerlendirmesine olanak sağlamıştır. Belirli bir değerlendirme yönteminin tek başına kullanılmasının, platformlar arasındaki oyuncu deneyimlerini ölçmek adına kesin sonuçlar veremediği belirlenmiştir. Sonuç olarak, oyuncu deneyimlerinde kapsayıcı analizlerin yürütülmesi için multimodal bir yaklaşımın gerekliliği belirtilmiştir. Ayrıca, tez süresince yürütülen her çalışma, platformlar arasında oyuncu deneyiminin bütünsel bir bakış açısıyla değerlendirilmesine izin veren belirli öğeleri ve özellikleri sunmuştur. Böylece kullanılan yöntemlerin veri üçlemesi yoluyla incelenmesi, platformlar arasındaki farklılıkların nedenlerini tanımlamak ve bulguların çapraz doğrulamasını yapabilmek için yararlı olmuştur. Oyuncu testleri ile birlikte farklı metodolojik yaklaşımların bir arada uygulanmasının oyuncu deneyimini ayrıntılı bir şekilde analiz etmek için etkili bir yöntem olduğu bulunmuştur.

Sonuç olarak, önerilen yöntem setinin uygulanabilirliği, oyun geliştirme sürecinin aşamaları aracılığıyla incelenmiş, platform farklılıklarını analiz etmek için kapsayıcı bir model sunulmuştur. Araştırmacı ve geliştiricileri daha iyi yönlendirebilmek adına, ‘Casual’ türdeki bir oyunun üretim aşamaları ele alınmış ve prosedürler ile kullanılması önerilen yöntemler ilişkilendirilmiştir. Bu bulgular, alanın daha da geliştirilmesi için farklı oyun türlerini veya platformlarını değerlendirmek için ek yöntem setleri geliştirmek için kullanılabilir. Önerilen bu yöntem yaklaşımı çerçevesi ve bu yöntemlerde sunulan alt öğelerin uygulanmasıyla, platformlar arasındaki oyuncu deneyimindeki farklılık daha iyi bir şekilde anlaşılabilecektir.

# 1. INTRODUCTION

## 1.1 Motivation & Purpose of the Thesis

Back in 1950's, commercial use of computers became available but came with a major technological barrier for the general population. For a time, only major corporations and universities had access to computers (Preece et al. 1994). In time, because of the technological advancement, computers became cheaper, faster and more reliable. These advances and novelties not only provided faster and advanced hardware capabilities, but also changed our perception towards them. Although these technological advancements provide vast and unexplored domains of potential uses, the need for a high-quality user experience has emerged and increased exponentially. To eliminate this gap between users and increase the utilization of computers, an effective interaction design was necessary. The field of Human Computer Interaction (HCI) has constituted to this need. The purpose of the HCI field is defined as; "*designing a computer system that supports people so that they can carry out their activities productively and safely*" (Preece et al. 1994). In other words, HCI field has the notion of making computers systems, applications and other software approachable.

Digital games share a similar origin. Until the late 1970's computer games were to be played among some lucky minority who had access to a computer. Even though the games developed were primitive in nature, they represented a cornerstone regarding the rise of popularity towards computer games. As the developments in information technologies advance, games remain to be one of the most challenging and popular applications among them. Because of their unique property of pushing the conventional use of technology, games are considered to be one of the main reasons for these novel technologies come in to our homes in the first place (Mayra, 2008). Computer games have also become one of the major components of entertainment industry in the last three decades (Korhonen, Montola & Arrasvuori, 2009). Now that both young and old people play games in their leisure time, games have become one of the most profitable market in

entertainment industry (Sanchez, Vela, Simarro & Padilla-Zea, 2012). Although the market for games is growing, only a few of them achieve success in terms of earnings. According to common marketing knowledge, the success of a product or a brand depends on its capabilities of satisfying the customers' needs and expectations (Kotler, 1993). This perspective may conceive to apply for games as well. Many studies indicated that the experience is much more crucial factor compared to a systems hardware and software capabilities. Tsekleves et al. (2009) mention that the needs and expectations of the user interaction yields a crucial role beyond hardware and software capabilities of a system. Therefore the users' expectations and habits are as important as the novel capabilities of technologies in this context.

Similar to the exponential increase of importance of HCI in productivity software and other applications, digital games have recently became a subject of research in terms of their interaction capabilities and the experience they are so good to deliver. In fact games as we know originated from the field of HCI. One of the first games developed on a computer come to exist when Alexander Douglas employed it for his Ph.D. thesis on human-computer interaction (1952). Ultimately games were meant to offer enjoyable and fun interactive experiences. Moreover, the notion of a game being enjoyable and fun refer to the experiences and perceptions of the players (Preece, Rogers, & Sharp, 2007). Since the main success factor for a game is considered to be the ability of delivering a satisfying experience to players, for the last decade, researchers realized the importance of user-centered approach and the utilization of the concept of user experience in the field of games (Bernhaupt, 2010; Nacke, Drachen & Göbel, 2010; Takatalo et al. 2010). Even so, the user-centered design principles are not at the same level as they are for productivity software and other electronic applications. Unlike other applications, games excel at offering enjoyable challenges and engagement without the necessity to be productive. Consequently the conventional usability approaches embedded in the HCI studies could not be applied to games directly. Sánchez et al., suggested that analyzing the quality of a game via usability evaluation is not sufficient (Sanchez, Zea & Gutiérrez, 2009). In the realm of games, it is necessary to go beyond the notion of interface usability to understand and evaluate other aspects of games such as game mechanics, story and gameplay

(Desurvire et al. 2004). Several researchers have studied methods for evaluating games by adopting methods from the usability field like conducting usability tests or heuristic evaluations (Federoff 2002; Desurvire et al. 2004; Röcker & Haar 2006; Korhonen & Koivisto 2006, 2007; Pinelle & Gutwin 2008; Schaffer 2007; Pinelle et al. 2009; Jegers 2008; Bernhaupt et al. 2007, 2008) yet never reached a consensus in detail and in quality of description. Because of the need for a novel perspective towards evaluating games, the term player experience (PX) was adopted by researchers in the field instead of user experience (UX), hence the term 'usability' was transformed to 'playability'. The context of playability goes much deeper than usability since it is not only related to the fun or enjoyment of the games.

However, there is still a lack of understanding on how the interactive capabilities of games relate to the topics of user experience such as usability and design (Sutcliffe & Hart, 2017). In this study, I will focus on both player experience and playability of games for evaluating the issues which define the gaming experience.

Today, platform differences between games also play a crucial role in terms of player experience. From conventional interaction methods of keyboard and mouse to the most advanced gesture-based interactivity capabilities, and from old Cathode Ray Tube (CRT) monitors to virtual reality (VR) headsets, it seems the technological advancements are never to cease. Accordingly, games are now played through various platforms including mobile phones, tablets, desktop computers and consoles. For instance, in recent years, more players have started to play games on their mobile phones over other platforms (Soomro, Ahmad & Sulaiman, 2013). Regarding the global games industry in 2017, from all of the global sales, the console platform took a share of %29 while mobile platform took %43 and PC platform took %28 ("GamesIndustry.biz Presents. The Year in Numbers", 2017). This diversity and shifting of platforms through sales indicate that the utilization of platforms for gaming has a crucial impact for game designers and the industry itself. Thus today it is crucial for the gaming industry to adopt a systematic approach for understanding and evaluating player experience on various gaming platforms (Zhu et al., 2017). Moreover, majority of the studies focus on content instead of gaming

platforms. Lee & Peng (2006) indicated that most of the existing literature usually focuses on the effects of the context and neglects the medium and its form. Yet understanding the form of the media and content (such as size, audio, motion) is as important factor for determining the effects of media. They added that:

> "*Studies on the main effects of computer games' form factors and possible interaction with content types are needed in order to get fuller understanding of game effects towards the gamers*" (Lee & Peng, 2006. p. 340).

Although there is a need for a multi-modal approach for analyzing and determining player experience (Poels, Kort & Ijsselsteijn, 2012), there are only a scarce number of studies within the literature that investigate player experience on different platforms, often only focusing on difference of control mechanisms (e.g. Suhonen & Vaataja 2010; Zaman et al. 2010; Gerling, Klauser & Niesenhaus, 2011). Understanding differences in player experience between different platforms inclusively will not only help identify the ways of player interactions but also help game designers and developers to create and evaluate the games that would meet the expectations of players (Nordin, Denisova & Cairns, 2014).

This dissertation is fundamentally a multidisciplinary work combining HCI research and game research. It was aimed to analyze and inspect various playability evaluation methods, and map out which method could be used best for evaluating certain aspects of digital games in the context of platform differences. In this study, instead of proposing a new methodology, I intended to utilize existing methods in a multi-modal combination to assess if and how they are effective in analyzing different player experiences between platforms. Game research and user experience research defined the domain of this study. In order to pave a way for analyzing player experience differences between platforms, it is essential to understand the characteristics of various game experience evaluation methods. The present research addresses issues with these various methods for evaluating playability of games, including analytical methods utilizing a usability-based evaluation, an empirical method with surveys, heuristic evaluation and user-testing.

My motivation for the study is also to provide a methodological set for player experience evaluators for future research. Because of this reason, my approach in this dissertation is a pragmatic one. Additionally, instead of deeply reviewing and getting involved with the debates about the definitions, facets and terms defined in the literature, I will focus on utilizing the existing research and evaluation methods to provide a practical guideline for game researchers and designers.

## 1.2 Background Review

### 1.2.1 Video Games

Technological advancements brought novelties such as diverse utilization options but also entailed the evolution of conventional technological devices which allowed new and user-oriented way of interactions to appear. The transition from keyboards to touch-screen interactions emphasized the importance of user-oriented interaction design. Following this notion, one of the most prominent area in technology which we can observe the advancement in both interaction and technology is the field of video games (now typically referred as ‘games’). As Mayra (2008) puts it:

“*Very few software applications are capable of delivering similar experience of genuine and rewarding interaction with the system as digital games do. It can be said that interactivity is what games are and what they do, at the very core of gameplay*” (Mayra, 2008).

For the last decade, our goals and means of using the technology in our everyday lives changed drastically. The habits and expectations regarding the use of technology and retrieving information have evolved accordingly (Seager et al., 2007; Tsekleves et al., 2009). This also brought games into our daily lives. We all probably have already know someone who plays video games on a regular basis either on a gaming console, PC at home or a smart phone on the go. The commercial success of games are not to be neglected as an annual sale of 116 billion Dollars are made from the industry in 2017 (“GamesIndustry.biz Presents. The Year in Numbers”, 2017). It is apparent that video

games industry is capable to compete with other entertainment industries such as cinema or music.

In order to define games, Esposito (2005), provided a definition for a video game: "*A video game is a game which we play thanks to an audiovisual apparatus and which can be based on a story*". According to that definition, the game in our era is consisted of; gameplay, audiovisual gaming platform and game story. Esposito's definition is derived from former literature on games and play such as (Huizinga 1950; Zimmermann 2004). In order to avoid any misunderstanding, in this study, I consider the term games as an umbrella including all forms of digital games independent of their platform.

Ancestor to computer games, electronic games are in our lives for over 60 years (Kent 2001). Regardless, video games as we understand were not available to general public until late 70's. The first game was developed at 1940's, such as the 'cathode-ray amusement device' (Goldsmith et al., 1948). The game simply simulated a missile being fired to a specific target via manipulating several knobs to adjust speed and curve of the missile. The game was played on a CRT television. Only a dot was displayed on the screen of the television but different elements were graphically placed on transparent printed layers which later applied on the screen to simulate a radar-like environment. The first video game as we understand, *Noughts and Crosses* or *OXO*, was developed on a computer as a part of the Ph. D. thesis on HCI from Alexander Douglas in 1952. It was a game that simulated the *tic-tac-toe*. The game was played by putting inputs to a rotating phone dial. The nine cell game board could be seen through a tiny 35 x 16 pixel screen. However, the game was capable to run on one of the most advanced computer back at that time, and general population did not have the chance to experience it. Later, in 1958, a game named *Tennis for Two* is developed by William Higginbotham. The game employed a small oscilloscope and two controllers rotating the trajectory of the tennis ball and its rebounding angle. The game was created as an entertainment tool and was successful to that end. Similar to *OXO*, *Tennis for Two* is developed in a laboratory, restricting its experience to a selected few. The first video game to attract some attention from a population is *Spacewar*. The game was developed at the Massachusetts Institute of Technology (MIT) in 1961 by a small group of students led by Stephen Russell. The game

was a combat style game in which two players were to attack each other avoiding the gravity well in the center of the play area on the middle of the screen. It also employed a bigger display compared to former games and received several updated versions featuring improvements such as better physics, star constellations as background image and button based controllers. After those improvements, the game received a quick popularity which made the game being shipped with the computer system as a demo application. Among the students whom developed *Spacewar*, Nolan Bushnell and his friend Ted Dabney developed the first commercial video game in 1971. The game *Computer Space*, was developed as a single-player version of *Spacewar* game. The goal of the game was to hit the computer generated moving saucer using the provided controllers while trying to avoid incoming fire. The game did not have the mechanics of "player lives" and instead the game sessions were limited to 100 seconds. The game also provided an additional 'hyperspace' mode which extended the play time 99 seconds more where the graphics turned from white-on-black to black-on-white.

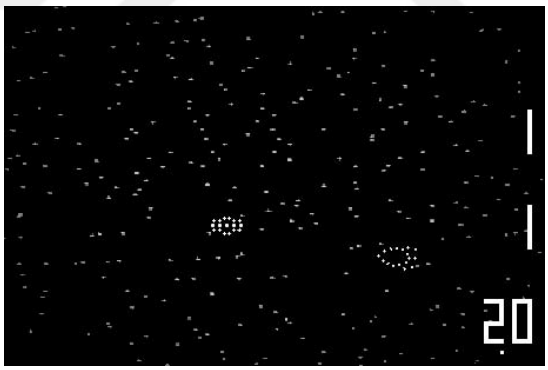



**Figure 1.1** Computer Space, the First Commercial Coin-Operated Game Ever Produced (Nutting Associates, 1971)

This early commercial attempt of Nutting Associates failed commercially simply because they did not understand the differences between various player groups, their capabilities and cultures. The game was playable by the MIT or Stanford students yet it was found too alien and complex for the general population.

While Nolan Bushnell and his associates were trying to bring computer games to general population, engineer Ralph Baer was the first person to bring video games to people's homes. He developed to first prototype of a system which works with a television set in 1967. It was a two-player ping pong game, *Table Tennis*, which the players were able to play on a regular television screen. Through several iterations, the system became capable of providing various game types. The players just needed to change a cartridge to begin a different style of a game while the system was connected to the television. The system was named *Odyssey Home Entertainment System* and it was a revolution and the beginning of console gaming. Thought limited, several peripherals could be attached to the console. In 1971, TV manufacturer Magnavox receives the rights of the license for the gaming system. Finally in 1972, the first gaming console, the *Magnavox Odyssey 1 TL200* arrived to the market. The console unintentionally led to the development of one of the most influential and successful game back in 70's. Even though Nolan Bushnell, founder of Atari and former developer of *Spacewar,* only saw the game *Table Tennis* in a *Magnavox Odyssey* demonstration in 1972, it was an important inspiration for him. Atari's arcade game *Pong* (1972) brought a huge and immediate success yet it was not a home console but an arcade game machine. The game was played more like a squash than ping-pong and it had become a game of angles which players were able to devise strategies banking shots against the walls on the screen (Kent, 2001).

**Figure 1.2** Spacewar on the Computer History Museum's PDP-1. Creative Commons/Kenneth Lu, the first 2D top-down shooter game

While explaining the history of video games, it is important to mention games that pushed the boundaries and influenced the whole gaming scene. After the success of *Pong*, games started to become a business in itself and inspired many engineers and designers to create various examples and genres of games.

One of the team members of Atari was a now famously known developer, young Steve Jobs who with his friend Steve Wozniak created the game *Breakout* in 1976 for Atari platform. The game was important for several aspects, including but not solely the capacity to revive the ball and paddle gameplay after *Pong.* Because of its minimalistic architecture in software, the game required a minimal amount of chips on the circuit board which was a crucial advantage and achievement at that time. The game was successful and Atari managed to sell tens of thousands of cabinets worldwide. This experience from the development for the game designers were so important that they facilitated many of the principals and perspectives (such as the efficient use of chips) from the game architecture to develop future Apple computers.

The following year, Atari released the game *Night Driver* in which the players were to experience the first game that provided a first-person point of view. The game had a black background (hence the night) and white road-side reflectors were indicating the road to follow. The small white rectangle at the bottom of the screen symbolized the car and the players were able to control it via realistic car controls such as gas pedal and steering wheel, further enhancing the immersion.

In 1979, Atari again published a game called *Asteroids* in which the player was to eliminate floating asteroid like vector graphics by shooting at them while rotating and moving around. The game was important and influential because it provided an increasing level of challenge through gameplay. The more player was to shoot asteroid, the smaller and faster moving parts they created. Sharp vector graphics and simple but effective sound design also contributed to the success. At the time when games were using plastic layers to colorize graphics, in 1979 *Galaxian* by Namco provided colorful sprites and animations as well as moving animated background. Even though it was similar to the famous game

*Space Invaders*, *Galaxian* had enemies which were capable of performing complex alien movements.

Another successful game from Atari, *Star Raiders*, was so immersive, players were playing the game day and night. The game was published in 1979 and was successful because of several reasons, the most important being the story of the game. Because of its success, *Star Raiders* ported to Atari consoles later in 1982. In 1979, *Zork* was an important landmark for interactive text adventure games. The game reached success not only because of its engaging storytelling (and humor), but also for the quality of its text parser system which had an advanced syntax architecture and was able to create a natural interaction with the player. The game was able to receive commands in the form of prepositions and conjunctions like "*slay thief with knife*".

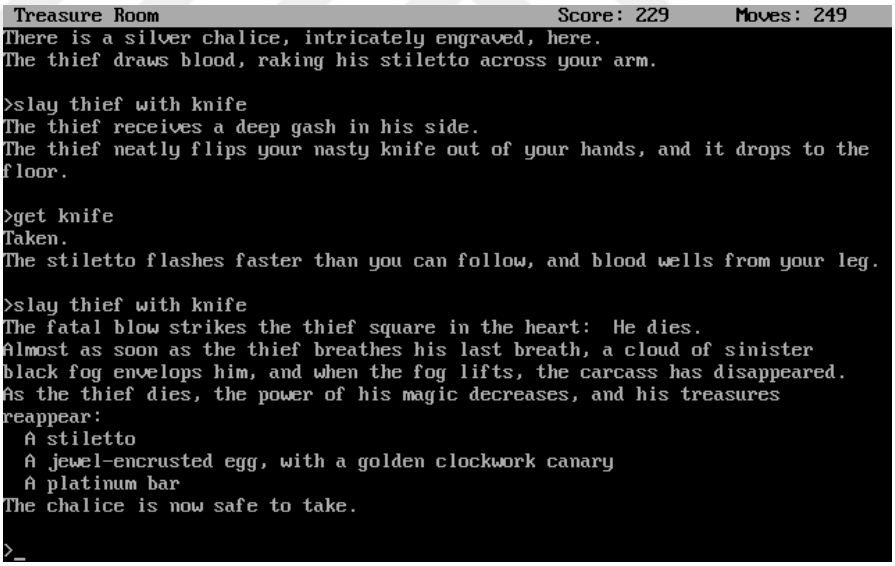



**Figure 1.3** Screenshot from the Text-Based Adventure Game Zork

*Tailgunner* was the first game that utilized 3D sharp vector graphics and animate shapes through few polygons in 1979. The game was developed by Larry Rosenthal, a former MIT student who worked on the game *Spacewar*. The player was put in the shoes of a ship commander as nimble enemy ships were approaching to your vessel with realistic

animations. The game was influential mostly because of its graphics and use of polygons for forming animations.

A year later, Williams Electronics published *Defender*, the first game in which players were to strategize their moves in order to win. It was also the first 2D side-scrolling game. With two joysticks and 5 buttons, the game was complex and provided enough challenge to players to master them. In 1980, *Flight Simulator* was released. As its name suggests, the game was about flying an aircraft combining both civilian and military flight maneuvers. The game offered two different game modes, one simulating a normal civil flight and one offering World War I British Ace game. The game also provided fast feedback regarding the cockpit animations. The following year was crucial for modern RPG game genre since *Ultima* was released on the Apple II in which the story was vast and immersive. The game then ported to other platforms (Commodore 64 and Atari) with the name *Ultima I: the First Age of Darkness.* The avatar of the player had 6 main attributes (strength, stamina, charisma, agility, wisdom & intelligence) which the player could increase his/her power by completing adventures. It was also the first commercial game to offer vast areas to the player to explore and advance. In 1982, Activision published *Pitfall*, a famous game for its graphics and complexity. The game also was important as it shed light to the capabilities of consoles by implementing complex graphics and game mechanics.

When observed, each of these early successful examples of games all benefitted from providing various novel game experiences that are still being considered as crucial under the umbrella of game experience research.

**Table 1.1** Early Influential Games and Their Featured Attributes

| ***Games*** | **Featured Attribute** |
|---|---|
| *Breakout (1976)* | Efficient use of hardware |
| *Night Driver (1977)* | First-person point of view |
| *Asteroids (1979)* | Exponential increase of challenge |
| *Galaxian (1979)* | Colorful animations and sprites & complex enemy movement |
| *Star Raiders (1979)* | Immersive story |
| *Zork (1979)* | Immersive story & advanced text interaction |
| *Tailgunner (1979)* | 3D vector graphics and realistic animations |
| *Defender (1980)* | Challenging gameplay |
| *Flight Simulator (1980)* | Fast feedback & different modes of games |
| *Ultima (1981)* | Explorable game area & immersive story |
| *Pitfall (1982)* | Advanced graphics and complex game mechanics |

### 1.2.2 Gaming Platforms

To understand the exponential success and popularity of video games and the increasing importance of player experience, it is crucial to understand differences between gaming platforms. Since games are designed in accordance with the interaction capabilities that are provided via the platform, it is necessary to inspect gaming platforms and their relation with games. Although Gerling et al. (2011) focused on how platform differences mainly differentiated the player experience in terms of controls, they indicated that the difference in platforms is not only important for the performance but also for personal experience in games (Gerling et al. 2011). Today, people are playing games on several devices and sometimes use more than one platform for gaming. Similar to the history of video games, this was not the case until late 90s.

The success of the first home console *Magnavox Odyssey* was followed by *Atari VCS/2600* in 1977, in which paddles and eight-direction joysticks were provided to allow different types of games. Atari in turn led to many success stories of 1980s, such as *Intellivision* (1980), *Colecovision* (1982) and *Nintendo Entertainment System* (NES; 1985). Later in 1994 Sony Entertainment released its first gaming console, *PlayStation* just after Nintendo's hand held gaming device *Game Boy*. Historically, both platforms were so successful that they broke records on global sales. *PlayStation*, have crossed the 100 million sold units line, and Nintendo's *Game Boy* series, sold 120 million units until 2006. Sony Entertainment and Nintendo still hold the leadership among the best-selling consoles as *PlayStation 2*, an iteration of the original *PlayStation* hold the most sold units record while *Nintendo DS* has the second place until 2017[*]

Naturally, arcade machines and consoles were not the only means to experience games. *IBM PC* was released in 1981. It was based on an Intel 16-bit chip and lacked any sound or advanced graphic capabilities. Moreover, PC's were also very expensive and seen as machines meant for work and not play. To overcome this price barrier and bring computers to homes, *Commodore* released the first 'friendly computer', Commodore *VIC-20* to general population, which was something that was completely unheard before. Because they were very low cost and popular, several games were developed for the system. Additionally, Commodore decided to emphasize the marketing strategy of *VIC-20* directly in competition with gaming consoles. *VIC-20* later become the first computer that is sold more than one million units. In 1983, Commodore was the first computer manufacturing company to report $1 billion in sales (Dillon, 2011). Later on, with the advancement of technology and more powerful 8-bit processors, *Commodore 64* started a new era by becoming a best-selling computer until the model was discontinued in 1993. *Commodore 64*, the best-selling home computer was released in 1982.

[*] Retrieved from *Gamespot.com*, Created by Jimmy Thang on October 31, 2017


**Figure 1.4** Commodore 64 & Nintendo Game Boy

PC architecture was fundamentally different than of consoles since it was backwards compatible. The architecture also provided third party companies the opportunity to produce additional hardware such as sound cards or graphic cards which in return revolutionized the PC as a gaming platform in 1990s (Mayra, 2008). For instance *Creative Technologies* was founded in 1981 in Singapore first released their sound card specifically built for PC in 1989 while other brands also started to emerge producing graphical cards. ATI technologies released the first ever graphics card, *Match8*, which could be integrated with both IBM and Commodore in 1991. Other now well-known companies such as Nvidia or Matrox followed this trend and released improved graphic processor cards. More power presented more possibilities for PCs, naturally including gaming.

Exploring these new capabilities, software companies often focused on the utilization of 3D graphics. ID software which was founded in 1991 released the first 3D PC game, *Hovertank 3D* in the same year. Although the game was seen as a milestone, their third iteration, *Wolfenstein 3D* was the real game changer. Followed by groundbreaking advancements such as online and LAN multiplayer capabilities, games such as *Doom*, *Starcraft* and *Quake* generated an extreme popularity to PC platforms. By the mid-90s the year of internet had become a new frontier for business. Meanwhile *Windows 95* was released for PC which caused PCs to be considered synonymous with it. With *Windows 95*'s accessibility and the lure of internet, late 90s is when PC's were mainstream among gamers alike.


**Figure 1.5** Screenshot from the Game Wolfenstein 3D

Although, following the huge popularity in the late 90s, PC gaming industry had a major crisis which lasted until early 2000s. PC companies of the past two decades such as Westwood or Origin have been either closed or bought by bigger companies such as EA or Activision. This limited the contextual development of games in PC platform. Next to the ever growing popularity of consoles, PC platform definitely lost some power. Moreover, in the hardware side, a similar crisis was happening at the same time. Once famous manufacturer 3Dfx suffering from bad management decisions bought by Nvidia, leaving the graphics card industry as the duopoly between ATI and Nvidia as it still stands today.

In the second half of the 2000s, giant software companies such as EA, Activision and Ubisoft as well as hardware companies such as Nintendo, Sony and Microsoft have dominated the industry of games. This oligopoly have led the productions to stagnate. In order to cope with this, companies have started to develop games for every possible platform.

Places where the game is played significantly changed with mobile technologies. Mobile platforms offered the chance to play games on subways or other public spaces outside of homes and offices. Mobile devices not only blurred the line between everyday life and private space but also allowed various applications to be used in public venues and even educational purposes. Kopomaa (2000) indicated that mobile media presented a

novel ‘third’ space rather than private and public space. Mobile games also are considered to have the same meaning with casual games which emphasize games being simple, plain and commonly accepted (Kultima, 2009).

Following the advancements, mobile games obviously could be traced back to first mobile hand-held devices but first game to be released on a mobile phone as we understand was the *Snake* which first appeared in 1997 on the Nokia 6610. Provided with the capability of a single colored screen and phone dials, players were able to interact with the game through pushing dial button for directing the snake, trying to collect as higher score as possible. To this day, the game is seen as one of the most influential mobile game ever developed. Surprisingly, a variant of the *Snake* for the Nokia 6110 was also the first two-player game for mobile phones, utilizing the infra-red capability of the phone. Since the power of mobile phones in terms of hardware cannot compete with other gaming devices, games developed for the mobile platform follow-up previously developed game adaptations for a time. For instance the game *Space Invaders*, originally developed in 1978 was adapted to mobile platform in 2002. Recognizing the potential for mobile gaming, Nokia released their game oriented phone *N-Gage* in 2003 to attract gamers away from Nintendo’s *Game Boys* by including mobile phone functionality. But the most game-changer device was released in 2007 by Apple. *iPhone* was capable of various multimedia capabilities including a touch reactive screen, allowing much more natural interaction possibility rather than static buttons embedded on previous mobile devices. Following its release and the launch of *App Store*, many very successful mobile games were and continue to be released such as *Angry Birds* (2009), *Plants vs Zombies* (2010), *Fruit Ninja* (2010) and *Temple Run* (2011). Making use of touchscreen capabilities, developers keep on producing novel interactions and games for mobile phones until today.


**Figure 1.6** Snake Game on Nokia 6610 & Fruit Ninja Game on 6th Generation iPhone

Even though in early days, games were developed for a single platform, in our era, which the game industry keeps on growing and expanding throughout various platforms, game publishers no longer can afford to neglect cross-platform development of their games. Instead they often mandate their developers to produce the game for more than one platform simultaneously. Because of this competitive scene in the industry, it is common to witness games being published with several brandings representing the available platforms for them to be played on: “PlayStation 2, Xbox, and GameCube” (Reimer, 2005).

### 1.2.3 Player Experience and Playability

#### 1.2.3.1 Player Experience

As the game industry advances, so does the research on games. User research on games seems to uncover important aspects on gaming experience. The concept of user experience that has emerged in recent years in human computer interaction studies has also guided the analysis and evaluation of games (Bernhaupt, 2010; Nacke, Drachen & Göbel, 2010; Takatalo, Hakkinen, Kaistinen & Nyman, 2010). Since playing games fundamentally mean to interact with a digital interface, studies related to game experience necessitate to take into account both characteristics of gameplay and general issues related to human computer interaction. Similar to the relation between the concept of 'user experience' and usability, gameplay and playability are considered to be parts of the more inclusive concept of 'player experience' (Figure 1). Nacke et al. (2009), argue that player experience is directed towards players whereas the playability is the evaluative process directed towards games. Additionally they indicate that playability methods aim to evaluate the gaming system including the games while player experience evaluation methods aim to evaluate the players to improve the gaming experience. Since the term 'player experience' refer to the overall experience delivered through a game, it was indicated that "Good playability of a game should be a perquisite for evaluating game experience" (Nacke et al. 2009).

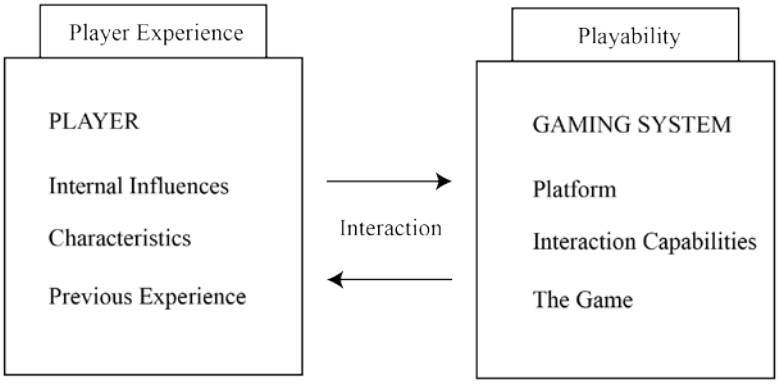



**Figure 1.7** Contextual Model of the Relation between Player Experience and Playability

Hassenzahl & Tractinsky (2006) indicated that the user experience is facilitated by usability and it includes two perspectives respectively; the *process* and the *product*. Similar to this notion, player experience is facilitated by the playability of the game.

Sanchez et al. (2009) view playability as user experience in games and defined the term as a set of properties defined to identify player experience of the gaming system which has the main purpose of providing entertainment and enjoyment by being credible and satisfying for the player (Sanchez et al. 2009). Additionally they indicated that a productivity application/software such as a word processor is developed with the aim to allow users execute a set of tasks in a predetermined context. But when it comes to games, the objective is to "*make the player feel good when playing i*t" (Sanchez et al. 2009).

Lazzaro (2008) also indicated the major differences between user experience in terms of productivity software and player experience in terms of games (Table 1.2).

**Table 1.3** Different goals for achieving positive experiences (Lazzaro, 2008)

| **UX: Productivity Software** | **PX: Games** |
|---|---|
| Task completion | Entertainment |
| Error elimination | Fun to overcome obstacles |
| External reward | Intrinsic reward |
| Outcome-based Rewards | Process is its own reward |
| Intuitive | New things to learn |
| Reduce workload | Increase workload |
| Assumes technology needs to be humanized | Assumes humans need to be challenged |

In addition, although it is also a type of user experience, it was indicated that player experience contains more complex features than any other user interaction with a system (Fierley & Engl, 2010). For analyzing games, well-known usability evaluation methods are considered to be inefficient when applied alone (Pagulayan et al. 2003; Korhonen, 2011). Hence, player experience as a concept which initially built-up with computer games, should be approached differently from user experience (Zhu, Zhao, Fang & Moser, 2017). Ultimately, it is only natural to observe games as being fundamentally different

than productivity software, which necessitates a different approach for evaluation rather than traditional usability research methods.

#### 1.2.3.2 Playability

Shaffer & Isbister listed three reasons for usability studies getting involved with the games; (1) developers and publishers are trying to reach out to broader audiences, (2) game development teams have grown, (3) proliferation of platforms (Shaffer & Isbister, 2008). They indicated that for these and many other reasons, game developers are turning to user research field for evaluating and advancing their work. In order to understand the importance of user research in the field of games, we need to define fundamentals behind usability research and how it is different from playability.

As the development and the increase in accessibility of computer technologies advance over time, it was necessary to aim for a general population and not only to people who have technical knowledge (Hartson & Pyla, 2012). The concept of usability is based on the logic of designing functions and processes offered by computer technologies (including games) not for people with a certain experience or ability, but for a large user population. Hence, usability has the notion of improving the user experience. According to ISO standard (ISO 9241-11, 1998), usability is defined as "*effectiveness*", "*efficiency*", and "*user satisfaction*" in a specified context of use. As Nielsen indicates; effectiveness and efficiency related with the user's ability and performance for completing tasks with the product at hand, and user satisfaction is a result of succeeding in task completion. A 'usable' product is generally easy to learn and use (Nielsen, 1994). With this conventional point of view towards usability, the researcher is considered to be only interested with a specific target group of people and task oriented goals to analyze if the product/software meets expectations of practical concerns. Common practice is to employ usability methods ranging from user observations, controlled user studies and/or inspection techniques (Barkhuus & Rode, 2007; Nielsen 1994; Dix et al. 1997).

Nielsen also mentions five components of usability (Nielsen, 1994). These are *learnability*, *efficiency*, *memorability*, *error-free use*, and *satisfaction*. He explained these aspects as:

> *“ 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. Further, catastrophic errors must not occur.*
>
> *Satisfaction: The system should be pleasant to use, so that users are subjectively satisfied when using it; they like it”* (Nielsen, 1994. p. 26).

Usability is commonly measured by having a sample group of test users which are selected to be as representatives of the target users. These participants use the system following pre-defined set of tasks. Also, usability can be measured in the field. The crucial point of these evaluations are, as Nielsen stated; “*is that usability is measured relative to certain users and certain tasks*” (Nielsen, 1994).

In accordance with the definition, in the last decade, the context of the term usability has widened. Productivity software developers took more interest in usability methods, putting the factor of engagement in the method to analyze the overall experience. Isbister & Schaffer (2008) underline this merge of contexts in the definition by indicating; “*This leads to testing techniques that are concerned with qualities such as engagement, flow, and fun—qualities that bring user research closer to the primary concerns of game developers*” (Isbister & Schaffer, 2008). Several game researchers (Federoff, 2002; Johnson & Wiles, 2003; Desurvire, Caplan, & Toth, 2004; Korhonen & Koivisto, 2006) have also stated that traditional usability evaluations cover only some aspects and issues

related with games and game interfaces and ignore gameplay aspects completely. Therefore it is safe to indicate that these fundamental differences were the reasons for the term ‘playability’ rather than ‘usability’, come to exist.

Although there are crucial differences, playability is closely related to usability. It is commonly related with the aspects such as learnability, memorability and effectivity by various researchers (e.g. Nielsen, Usability Engineering, 1994). This indicates there are similarities as usability has in productivity software. There are several early studies trying to define playability (Fabricatore, Nussbaum, & Rosas, 2002; Järvinen, Heliö, & Mäyrä, 2002; Sánchez, et al. 2009). Egenfeldt-Nielsen et al. (2013) have defined that good playability refers to fun, challenge and ease of use. Usability Glossary (2002) defines that playability includes and affected by storyline, controls, pace, usability and possibly other aspects. It is probably the most developed definition for playability so far as it 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.”*

In several studies, researchers have used the term *game usability* when they are relating the term playability through usability (Fabricatore et al., 2002; Isbister & Schaffer, 2008; Pinelle et al., 2008a). Fabricatore et al. (2002) have defined playability as understanding and controlling gameplay.

As mentioned previously, to fulfill the necessities of usability such as satisfying the needs of the user and to enhance the overall user experience, there are various user evaluation methods. For instance task oriented user tests gather data from a sample of users interacting with an application. Users are observed during the tests while they are asked to follow certain tasks. In the end, the results are analyzed to indicate (if) any issues related to the system which might affect the user experience. Differently, in the field of

games, playtesting methods often make those observations while the sample group of players interact with the game without any pre-defined tasks. In the literature of game design, playtesting method is mentioned frequently for evaluation purposes (e.g. Fullerton, 2004; Rouse, 2001; Schell, 2008). Correlatively to user testing, playtesting often requires moderation and observation during play. Think aloud methodology is frequently employed during the tests, while surveys and/or interviews are administered after the gameplay.

Alternatively, applicable to both fields of productivity and games, heuristic evaluation offers the benefit of evaluating the interactive application during the design process and most importantly, do not require a task oriented inspection. This method is also considered as an inspection technique which allows evaluators to examine an interface using statements of usability (or playability in the field of games) principles (Nielsen, 1994). Since user centered testing is becoming widespread for the industry of games (Isbister & Schaffer, 2008; Pagulayan & Steury, 2004; Wixon & Pagulayan, 2008), several sets of heuristics have also been proposed for evaluating playability. For instance Federoff (2002) first proposed a set of heuristics derived from traditional usability heuristics. Later, Desurvire, Caplan & Toth (2004) also proposed a set of heuristics with four categories; gameplay, game story, game mechanics and game usability. Korhonen & Koivisto (2006) proposed '*playability heuristics*' for mobile games including three modules; gameplay, game usability and mobility. Desurvire et al. (2004) also conducted a study for comparing methods of playtesting and heuristics evaluation. They administered specific heuristics that they have prepared for the context of games. The results of the study indicated that both methods identified problems regarding the interface whereas the use of heuristics identified general interface issues. A much more detailed explanation of playability evaluation methods have been provided in the second chapter of this dissertation.

To summarize, it is understood that player experience included diverse attitudes and feelings of the players' while playability focus on the gaming system and the game design elements. More so, identical to the relation between 'user experience' and 'usability', 'player experience' is facilitated through 'playability'. The importance of games regarding

both the industry and academia keeps on growing, but the subject of evaluating player experience has been a complicated field of work because of the following reasons: (1) Literature on player experience and playability is not profound; (2) there is lack of common definition for playability and there is no consensus on the methods for evaluating playability; and (3) researchers often employ different and novel approaches and therefore offer different set of heuristics for evaluation.

### 1.3 Objectives and Research Question

Based on the background review presented in the previous section, it is possible to make some conclusions and define relevant research questions. This present study aims to add to the existing knowledge on player experience and playability by combining various evaluation methods from game research particularly on evaluating how different gaming platforms offer different player experiences. Since playability is directed towards the gaming systems and games, it is crucial to evaluate the playability differences between platforms. Earlier research has shown that various methods aiming to analyze and evaluate player experience and playability lacked the notion of highlighting differences of experience between platforms. By inspecting previous literature it is possible to make observations to some extent: Individually, evaluation methods used for explaining gaming experiences and differences in playability between platforms are insufficient. However, an inclusive scientific method will potentially be able to close this gap. It is possible to make a comprehensive evaluation of platform-specific experience differences through multi-modal player experience analysis.

Emerging from this topic, I will explore in this dissertation:

I. Whether there are significant differences in the gaming experience between the platforms and those differences can be explained by various playability evaluation methods covering the player experience for games,
II. If so, whether it is possible to identify the differences in player experience between platforms can be presented inclusively by a multimodal evaluation.

In my research, I will explore these research questions following a mixed methodology, combining knowledge from both the field of HCI and game research. I will propose a multi-modal approach utilizing playtesting, player experience surveys, and expert evaluation methods to provide an inclusive coverage for determining differences of player experience between platforms.

### 1.4 Contributions

This study presents two main contributions following and addressing various aspects of the proposed research questions:

A methodological literature review to identify potential contribution for the application of various playability evaluation methods for assessing player experience. It is anticipated that the comprehensive literature review will also provide basis for further developments on evaluation methods analyzing issues which game designers and researchers consider important when they speak of playable games.

Secondly, by employing a multi-modal approach, a series of experimental researches will provide evidence on how well these methods are applicable in analyzing differences in experience between platforms. Results of those experiments will indicate strengths and weaknesses of those methods when administered by themselves and in combination.

The tangible contribution of the study will be a ‘methodological kit’ to evaluate player experience for game researchers and designers. The proposed methodological set would provide an overarching framework to indicate which playability evaluation methods are viable for analyzing differences between platforms regarding the gaming experience. The research results help in understanding and differentiating efficiencies and effectiveness of various playability evaluation methods.

### 1.5 Structure

This first chapter introduces the motivation and background of the research, describing the origin of this study and the derivation of its objectives. Additionally the research questions and objectives are described together with the overview of the structure of the thesis.

The second chapter will present a summary on fundamentals of playability evaluation methods will first guide the reader to diverse methods used for evaluating games and their playability. This section will describe methodological analysis of playtesting and heuristics evaluations methods for assessing playability of games. Moreover the implications will be discussed.

In the third chapter, the exploration of player experience in different gaming platforms in which domain specific issues and analysis of platform differences will be discussed. Special emphasis was put into the analysis of conducting multi-modal evaluation for analyzing playability issues.

The experimental studies and their results will be discussed at the fourth, fifth and sixth chapters where the goals, methodologies, analysis, and results of the studies will be explained. These studies include playtests, surveys, interviews and expert reviews for analyzing platform differences in the context of player experience. To inspect various evaluation approaches three different experiments and evaluations on analyzing player experiences were conducted during the study. A mixed-method approach was employed during the studies in order to analyze player experiences holistically. Some evaluations were conducted with expert evaluation method while some utilized questionnaires, yet playtests were conducted for all these studies for inclusive analysis.

Lastly in the fifth chapter, results and conclusions derived from the studies and their role in the context of this dissertation will be explained. The dissertation will end with implications drawn from these studies that hopefully will provide a way for future research in analyzing playability of games.

# 2. FUNDAMENTALS OF PLAYABILITY EVALUATION

## 2.1 Methodological Analysis of Playtesting

### 2.1.1 General Approach

The game industry has been developing immensely although the methodology on evaluating player experience still lack a robust approach to evaluate the overall experience. User tests in productivity software proved to be very efficient in terms of analyzing a system and they are capable of providing in-depth feedbacks and insights. Similarly, playtests have the potential to present actionable insights and feedback from the players regarding the game design. Although user tests are effective for analyzing some aspects of a game, they often lack the possibility to gain knowledge about player's perception towards the game since playtests are usually focused on the player's behavior during the game. Despite playtests are usually focused on the player's behavior during the game, several researchers defined the term differently. Even though several studies utilized playtests for both validation and formation purposes, different evaluation methods are described in the following section. The methodologies and differences in procedures are summarized to be able to draw a framework for understanding how playtests were applied.

In their study, Davis et al. (2005) indicated that they have "*combined surveys with hands-on gameplay into a method*" which they called the playtest method. According to their definition, playtest is when the tests are conducted to obtain feedback about player experience from a specific game in a scientific manner. The goal for playtests is to obtain actionable feedback from the players to provide insights about the crucial elements of a game. They also mentioned that playtesting method can be invaluable especially if it was

used to assess to compare different player control schemes. In spite of various benefits of a playtest, it was also indicated that it may lack to evaluate important insights that may have been gathered from other usability studies.

Furthermore, Fullerton (2008) defined playtesting as not just gathering observations and feedback from the player but is a process which "*the designer performs through the entire design process to gain an insight into whether or not the game is achieving your player experience goals*" (Fullerton, 2008). As is seen from this definition, the main purpose of the playtesting is to achieve a positive player experience rather than only finding bugs, getting feedback or improving the interface. In the study, it was also indicated that the recruits for the playtest sessions should be chosen from players which represented the target group of the game being tested. It was also advised to isolate participants during the tests to prevent any social influence. Moreover, playtesting can be seen as a method for analyzing the player experience either early on or when large parts of the game are playable. Fullerton (2004) mentions that playability test(s) should be done in the early stages of the game design to allow the possibility of fundamental changes in game mechanics and dynamics. In contrast, Rouse argues, that "bringing playtesters in too early will only delay the game's progress" (2001).

Even though there are several different contexts and definitions that the term is used for playtesting, in methodological terms, most of the literature point towards a common way of application. First of all the playtests sessions include observations by default. Secondly, with some exceptions (for ex: Järvinen, 2005, Majgaard & Jessen, 2009), playtest sessions included interviews. A literature review has been conducted to assess methodological differences and similarities between playtests.

### 2.1.2 Playtesting Procedures

In order to analyze the literature regarding playtesting, a search was conducted with the keywords of 'play/testing' and 'playtesting'. A search in ACM Digital Library, IEEE, Springer, Taylor & Francis, Google Scholar and ISI databases was conducted. The result were then inspected by their abstracts for their relevance. This procedure is followed by a

full review of these findings. These articles were examined in detail to find out their main contributions to the research are, resulting in 14 relevant articles left in total. It should be noted that although the heuristics evaluation methods are going to be examined in the next section of the dissertation, some of the articles that are reviewed from the literature included playtests as well as heuristics. These multi-modal studies are mentioned in the next section. In this regard, it is crucial to underline that in this part of the chapter, the articles that focus on playtests are going to be examined.

Davis et al. (2005) have proposed a study in which they presented a description of playtest method and derived from this definition, proposed their procedure of playtesting. In the study, they conducted a case study for analyzing the effectiveness of the playtest which utilized both scientific surveys and supervised laboratory tests. Researchers have tested the Xbox game *Brute Force*. Although they have explained how to conduct a playtest and how it is similar to user tests, details regarding the actual evaluation procedure were absent in their study.

Järvinen (2005) followed a playtesting approach for improving a card game. In the study, the game was iteratively improved by playtesting without any systematic method such as interviews to get feedback from the players but opted for a free-form debrief from the participants. It is mentioned that this method of application nevertheless provided results complimenting the general observations. At first, the playtests sessions were conducted with 4 participants. Later on, after the iterations, playtest sessions continued via recruits from industry. As a result they presented a card game *'Gamegame Solitaire'* where individuals or teams could play the game with a pre-determined constraints to provide ideas for a design of a game.

Winn & Heeter (2007) have conducted a study in which conflicts in educational game design was inspected via playtesting. The study included three playtest sessions, each iteratively improving the game. 10 participants were recruited for the first playtest sessions in which each subject was tested one at a time while researchers observed players engagement followed by an interview. With the second prototype, researchers again conducted playtests with the same amount of participants. The updated game has been

tested for the third time again but this time with six participants. As the end of the study, they indicated that iterative prototypes and playtesting is crucial for designing a learning game. They concluded their research by indicating that rather than only helping designers, playtests could help "*resolving conflicts among pedagogy, content, and gameplay*" (Winn & Heeter, 2007).

Majgaard & Jessen (2009) have conducted the research for playtesting the use of analog and digital use of a playground by implementing a sort of a digital game into an actual playground. They conducted the investigations by solely observing the designated area of the place via unsupervised camera recordings. As a method of analysis, they preferred comparative analysis, evaluating differences between analog and digital play. For 14 days, they recorded the specified playground area and analyzed the sequences when the playground was occupied with children. The children were given no instructions on how to interact with the digital playground. As a result, the data they have gathered was based on the observations conducted via the recordings. To conclude, they indicated that analog play offers more variations compared to digital play and with the help of game designers, this lack of sufficient possibilities can be overcome.

Following year, Korhonen (2010) have researched the differences between expert evaluations and playtesting method. In the study, 6 participants were recruited and the test was conducted in a usability laboratory and one participant was playing the game at a time. Since the study was focused on analyzing mobile games, only participants who had mobile gaming experiences were chosen. In the study, Korhonen mentions that the playtesting sessions followed a standard user-testing procedure; including think aloud method, supervision and observation and a post-test interview. At the end of the study it is concluded that playability issues were much harder to indicate in playtests compared to expert evaluations and therefore two methods should be used as a complementary to each other.

El-Nasr et al. (2010) have conducted a study to understand and evaluate four different console cooperative games; Rock Band 2 (Electronic Arts, 2007), *Lego Star Wars* (Lucas Arts, 2007), *Kameo: The Elements of Power* (Microsoft Game Studios,

2005), and *Little Big Planet* (Sony, 2008). They were all playtested on a console with a total number of 60 participants. All the participants were invited to come as groups formed between 2 to 4 people. Players were interviewed for their general information at the beginning of the tests followed by 10 minute play sessions. Play sessions were recorded during the gameplay for further analysis. At the end each participant was interviewed a second time about their experiences and feedback regarding the games. In the conclusion they proposed a set of *Cooperative Performance Metrics* (CPMs) used for analysis of the cooperative games.

Mirza-babaei et al. (2011) have proposed a rather different approach on playtesting methodology. In their study they first categorized the overall playtesting approaches into four methods; *observation, think-aloud, heuristic evaluation* and *interviews & questionnaires*. To add on those categories they proposed the utilization of physiological measures. Before the tests, they gathered information on potential recruits. Afterwards, they recruited the participants according to their experience on games, preferences, console exposure and demographics. They have recruited six participants between the ages 20 to 31. They have conducted a series of experiments in a game user research laboratory with three experienced evaluators as supervisors. The games were played on a *PlayStation 3* gaming console. Participants were asked to play the first two levels of '*Call of Duty: Modern Warfare 2*' (2009) and '*Haze*' (2008). The biometric data was gathered during play sessions. In their study they utilized the biometric data and video recordings in order to pinpoint exact moments to construct a log which would then be discussed with the participant on an individual level. After the gameplay sessions, the logged recordings were shown and discussed with the participants. In the same study, they additionally used an observation based approach. Two evaluators analyzed the game footages to provide notes on usability and user experience related issues. Biometric data was not taken into consideration in this part of the study. In conclusion, they indicated that observation based methods would be effective for evaluating video game usability and game mechanic evaluation, as opposed to biometric analysis, which is better suited to the discovery of issues concerning gameplay.

Gerling & Masuch (2011) have studied the player experience of *Nintendo's Wii* platform with elderly participants. The article discussed the accessibility of commercially available video games for elderly players by conducting a focus group analysis. *Wii Sports* and *Wii Fit* games are chosen for the tests. During the play sessions, games were controlled by only the *Wii Remote* peripheral of the console. The average age for participants was 82, ranging from 67 to 91. Ten senior citizens were recruited for the playtests with no prior gaming experience. Because of the lack of gaming experience, researchers mentioned that participants had four gaming sessions in which they were able to test all the games. Each play session lasted for two hours and after the tests, focus group meetings were conducted. After discussions, follow-up interviews were made. Meanwhile, during the play sessions, participants were observed and supervised. Researchers indicated that the average perceived player experience in focus groups were positive yet observations during the playtests suggested otherwise. They also advised the use of user-centered evaluation approach for addressing specific requirements for specific audiences.

Despite being a conclusive study, Tan & Johnston (2011) have presented a research progress report of an automated playtesting process. They mentioned four different categories forming playtesting methods including; post-game reports, direct observations, In-game data hooks ad physiological measurements. Although among other methods, data hooks method seems unfamiliar, the term referred to game metrics at its core. In their research they planned on evaluating the tests via supervised learning system that maps physiological features to playtest questions. For the duration of playtests of each player, the presented system would continuously record important common expressions of the players such as expressions indicating happiness, sadness or anger. Nevertheless, the researchers noted that the proposed method was not an alternative to player-centric qualitative playtests but instead a supplementary measure.

Following up this study, Tan et al. (2012) published a study on evaluating facial expressions analysis to assess player experiences. In their study, researchers criticized the intrusive nature of playtests such as laboratory environments and proposed an automated facial expression gathering tool/method. In order to validate the feasibility of this approach, they conducted user studies. During the tests, a conventional playtest approach

was followed. 12 participants were recruited from various backgrounds and aged between 20 and 48. For the evaluation, two PC games were employed: A first-person shooter game *'Portal 2'*, and a trivia game *'Draw My Thing'*. After collecting demographics of the participants, researchers allowed them to play one of the games while they recorded physiological data. The researchers indicated that there were no tutorials regarding the tests and players were left to figure the games pout by themselves. After the tests Game Experience Questionnaires (GEQ) (Nacke, 2009; Ijsselsteijn, 2008) were administered yet researchers indicated that results from those questionnaires were not used for the purpose of the article. At the end of the tests sessions participants were asked if they noticed or felt intrusive about the recording equipment that were around during the play sessions. In conclusion, researchers indicated that the initial evaluation of the feasibility of the system have the potential to be a supportive measure for the player experience evaluations. The researchers noted that the tests were not to analyze the games at hand and/or the validity of the system being used but the effectiveness of the data gathered from the system.

Warren et al. (2011) have proposed a review on the application of usability and playtesting evaluation. They used the term usability as other researchers would use playtesting. In a similar fashion Warren et al. indicated that they have conducted a five phase usability test to evaluate the *'Chalk House'* literacy game. The usability test series that were conducted with the game have kept on over a period of 18 months prior to completing development of the product for widespread use. In phase 1, they conducted the tests in a laboratory with non-teachers and/or game designers. A supervisor first introduced the system to the participants. The participants were given task oriented instructions and asked to think-aloud into an audio recorder during the test. Tests are also recorded via camera. At the end of the phase, a follow-up interview was employed. In the second phase of the usability testing, researchers recruited end-users (teachers). Similar to phase one, usability evaluations were conducted. Differently in this phase, usability surveys and in-depth semi structured interviews were also employed. In the third phase, five sixth grade students were recruited as participants. Think-aloud approach and interviews are employed similar to previous phases. The fourth phase involved the same testing approach but within the confinements of an actual classroom. The fifth phase was described as the actual pilot study of the game, using three or more classrooms at once. In

the end, the researchers concluded that using at least one phase of the tests would be beneficial for any educational game. It was claimed as a mandatory step to implement actual users to the tests and the development process.

Quinn et al. (2013) presented a case study for describing the use of a mixed method approach to playtest a serious computer game. The study incorporated in-depth analysis of play session recordings, a survey and a follow-up focus group. According to the researchers there are two facets of playability as *informal* and *formal* protocols. The informal playtests were conducted via four participants from the internal playtesters of the inspected project. The participants were encouraged to indicate problems as well as productive criticism regarding the game. They were asked to take notes by themselves during play sessions. These tests were employed weekly for the iteration of the game project. By formal playtesting, the researchers referred to a more similar playtesting approach in which the participants are recruited from volunteers external to the project team. The process entailed following a playtest protocol which included observations, post-test interviews and focus group discussions. For this version of the playtests, multiple cycles of formal playtesting were conducted with a total number of 35 participants. The post-test interviews included questions to assess both behavioral responses as well as attitudinal responses which were quantified via Likert scale answers regarding some of the questions. Following the interviews, focus groups were used to discuss more detailed insights of the participants. The researchers indicated that focus groups were led with simple questions such as what parts of the game the players liked. Although researchers did not mention a conclusive result of the tests that they have conducted, they noted the article was aimed to provide a comprehensive method for playtesting by implementing various methods.

Zook et al. (2014), has based their study on a similar notion from Tan et al (2012) in terms of supporting playtests via automated measures yet they did not include any qualitative measures in the experiments. To conduct experiments on machine based playtests, researchers assigned specific parameters for specific actions. They empirically evaluated action learning by deploying two versions of an online game played on PC via web. Participants were asked to try play the game 10 times. Researchers then analyzed the

metrics received from the system and analyzed the results, making comparisons between player metrics and preferences such as A/B testing in the field of UX. As a result, they indicated that the automated system might provide useful for analyzing low-level design parameters and tuning but would definitely need a high-level approach which still dictates a human supervision.

Mirza-babaei et al. (2016) have conducted study on playtesting on a smaller scale feasible for indie studios. In their study the researchers aimed to understand and report the collaborations with six indie game studios by highlighting the value of playtesting process for indie developers and designers. Based on a previous research conducted by Mirza-babaei et al (2011), their approach similarly heavily relied on observation data from the playtests. The tests were conducted in a laboratory in which the play sessions were recorded. The tests reported to last about an hour divided for play sessions and post-interviews. The results regarding six different games from different developers were explained in their study. They presented eleven playtest cases studies on six different commercial games with 118 participants in total. Researchers noted that they were focused on only one area of each game (first time user experience) but needed more tests for evaluating other game elements such as challenge, or pacing of the game. They also highlighted that employing interviews were extremely effective due to their ease of deployment.

Cross references between playtesting methods that are followed throughout the literature are represented below (Table 2) with the notion of listing the available and most commonly used playtesting methodologies and procedures. According to this and the literature review, nearly all of playtesting sessions involved observations and play sessions. Moreover, interviews are the second most common method for playtesting. And lastly, the third most common method for conducting playtesting is recording the play sessions for later analysis by supervisors and/or researchers themselves.

**Table 2.1** Cross-referenced Playtesting Methods

| ***Playtesting Methods*** | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Game Metrics | | | | | | | x | | | | | | x | |
| Focus-group | | | | | | | | x | | | | x | | |
| Physiological Measurements | | | | | | | x | | x | x | | | | |
| Audio and/or Visual Recording | | | | x | | x | | | | | x | x | | |
| Interviews | | | x | | x | x | | x | | | x | x | | x |
| Free-form | | x | | | | | | | | | | | | |
| Observation | x | | x | x | x | x | x | x | x | x | x | x | | x |
| Surveys | x | | | | | | | | | x | x | x | | |
| ***Previous Playtesting Research*** | Davis et al (2005) | Järvinen (2005) | Winn & Heeter (2007) | Majgaard & Jessen (2009) | Korhonen (2010) | El-Nasr (2010) | Mirza-babaei (2011) | Gerling & Masuch (2001) | Tan & Johnston (2011) | Tan et al. (2012) | Warren et al. (2011) | Quinn et al. (2013) | Zook et al. (2014) | Mirza-babaei (2016) |

## 2.2 Methodological Analysis of Playability Heuristic Evaluations

### 2.2.1 General Approach

Among several playability evaluation methods, heuristic evaluation method is considered to be beneficial, not only for its capability of addressing issues, but also in terms of application. Heuristic evaluation is used mostly during the design process of games and do not require a task oriented test. It is defined as a technique that allows the researchers to conduct inspections using statements from usability principles (Nielsen, 1994). It is an inspection method used widely in usability research while game researchers recently started to study it. Korhonen indicates that heuristic evaluation is a method which is more effective at inspecting games rather than other methods because it does not require any task oriented research (Korhonen, 2010). Additionally, heuristic evaluation method offers to conduct tests in a fast and cheap manner since it necessitates a minimal number

of testers. So far, researchers studying the topic have not been able to propose a holistic set of heuristics which would be considered as common grounds for evaluating player experience and playability. Only a limited number of studies have tried to offer novel sets of possible heuristics related to playability of games (Malone, 1982; Federoff, 2002; Desurvire et al. 2004; Pinelle et al. 2008; Korhonen & Koivisto, 2006; Korhonen & Koivisto, 2007). Deciding on which heuristic to employ in a selected gaming platform is a trivial task. Because of this problem, some researchers have proposed different heuristics for specific platforms such as tabletop, computer, mobile, educational, social and web-based games (Korhonen & Koivisto, 2006; Korhonen & Koivisto, 2007; Köffel & Haller, 2008; Pinelle et al, 2009; Song & Lee, 2007; Tan et al, 2010; Suhonen & Vaataja, 2010; Liao & Shen, 2012). Moreover, the literature indicates that each heuristic method have their strengths and weaknesses.

In order to analyze the literature regarding the playability heuristics, a search was conducted with the keywords of '*game heuristics*', '*playability*', '*playability heuristics*', '*player experience*' and '*game heuristics*'. A search in ACM Digital Library, IEEE, Springer, Taylor & Francis, Google Scholar and ISI databases was conducted. After collecting the studies, a second examination over the abstracts have been done, leading to an elimination of irrelevant journal articles and proceedings. 44 articles were to remain as a result. These articles were examined in detail to find out their main contributions to the relevant research area. Because of this review process and the analysis conducted after it, some for the key differences for presenting heuristics for player experience have been observed. These key features regarding the differences were categorized in terms of what procedure was followed for the development of heuristics (such as the choice of medium or source for identifying heuristic items). In the end, four distinctive evaluation methods have been identified: Empirical evaluation, expert evaluation, inspection and mixed method approach. These evaluation methods and studies related to each category will be summarized and explained. Total number of 6 playtests sessions have been conducted for the study.

### 2.2.2 Heuristics Evaluation Procedures

In order to present categorization of playability heuristics, a number of studies have presented relevant articles which describe playability heuristics in a chronological manner (e.g. Jerzak & Rebelo, 2014). Although it has its benefits depending on the research, such as underlining the iteration process between evaluation methods and design approaches, it is not sufficient for analyzing differences between methods in terms of practical application. Hence, categorizing differences between heuristic evaluation methods in terms of procedures may hold a better perspective and potential for contributing the analysis process. Moreover, it made possible to indicate that there have been four main approaches in the field of player experience research. Among the eligible 44 articles, 12 articles were identified as fit to the empirical evaluation category while 17 articles fit in the expert evaluation. Additionally, 7 articles were eligible for inspection category and lastly 8 articles were observed to be fit for the mixed-methodology category.

As can be seen from these numbers, expert evaluations conducted by employing a set of proposed heuristics is the most common approach in the literature. This is followed by empirical evaluation method. Although many researchers advised that the most efficient and viable approach for a robust analysis is to employ a mixed-method approach, only 8 articles have used this method in practice (Figure 2). Similar in numbers, inspection method is often a focal point for criticism among researchers since it mostly gathers data from online reviews which possibly have various agendas rather than solely analyzing the games. Although I previously indicated that the chronological fashion for reviewing the literature is not beneficial in general, in the context of procedures, each article will explained under their specific category is listed in a chronological order to indicate the scientific progress in the field as well.

### 2.2.3 Empirical Evaluation

The first examined approach among various is the empirical evaluation method. This category includes researches conducted via user-tests, surveys, interviews,

observations and focus groups. The minimum sample group in this category in examined to be 10 participants. According to the literature review, 12 relevant articles chose to evaluate either the heuristics or specific games in their study via user testing rather than directly employing heuristics during research.

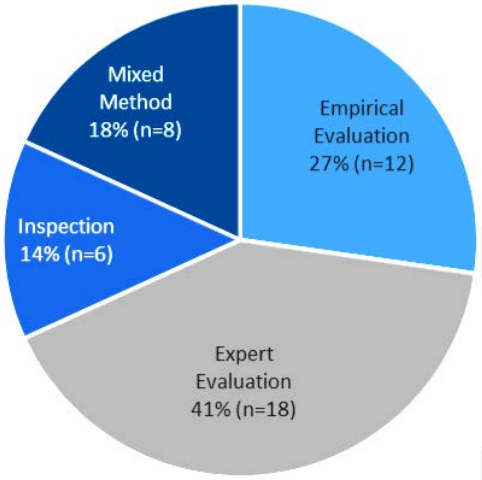



**Figure 2.1** Distribution of Different Methods Utilizing Heuristics Evaluation for Analysis

Malone (1982) have proposed the first set of heuristics for fostering games in the fields of learning and teaching. He proposed a set of heuristics for instructional / educational games and suggested three main heuristics to achieve an entertaining interfaces. In the study, a variant of three empirical tests were conducted to understand what gamers liked. A total number of 81 participants from elementary or second school students were recruited for the tests. The tests were conducted via three different games with 8 various versions for each game. In his study, he concluded with presenting three heuristic categories. These were '*challenge*', '*fantasy*' and '*curiosity'*. These three factors were broader frameworks including various aspects about enjoyable games. The category of '*challenge*' is related to the difficulty of a game related to the skills of the player and achieving goals (or generating them). '*Curiosity*' is more about motivation to learn and advance in the game as well as application of skills to accomplish different goals. In general Malone suggested an "*optimal environment*" in which the player should receive an amount of information but also these environments should evoke the feeling of

curiosity. Lastly, '*fantasy*' is about the narrative in which the game story, theme and how skills are used in the game is included.

After twenty years, Fabricatore et al. (2002) followed a different approach by presenting the heuristics first and testing them later. The heuristics model was prepared for guiding game designers and researchers for preparing better and more enjoyable games. It was a qualitative model of player preferences in action games that would evaluate and determine the playability in these games. 53 participants, between the ages of 20 – 30 have been recruited for the tests. The participants were asked to play 39 different games from different genres when tested. They were allowed to make comments along the tests in every step of the way. After play sessions, semi-structured interviews were conducted. By analyzing gameplay sessions, interviews and observation notes, they iteratively revised the proposed model. In the end of their study, they concluded with proposing a hierarchical model presented as determinants of the playability in games with a number of sub-categories in each. These indicated three main determinants were; '*entity*, '*scenario'*, and '*goals*'. The determinant of entity consists of four sub-categories; identity, energy, equipment and behavior. Scenario determinant consists; view, spontaneous changes, transitions and interactions with entities. Lastly, the determinant of goals include; complexity, linearity and interface. This research is considered as one of the first qualitative playability evaluation model present in the literature and therefore proposed a basis for the research field.

Getting inspired by the heuristic analysis approach, Röcker & Haar (2006) investigated if the present heuristics may well be used for evaluating pervasive games. The research was based on Desurvire's (2004) HEP heuristics. They conducted their analysis with a focus group of 10 participants, having varying ages between 32 – 28. The analyzed heuristics set was not shown to the participants during the tests since they mentioned it would jeopardize the objectiveness of the participants. A smart home environment with pervasive competencies have been chosen as the scenario for the participants in order to receive comments from them. After the interviews, participants were invited to join a focus group meeting and requested to take notes of their thoughts to cards that have been given to them. In the end, they were asked to conduct a card-sorting

exercise for ranking priorities and clearing up various categories regarding their comments. The researchers compared the heuristics with the proposed comments and indicated the need for additional heuristics to the set for it to be applicable for pervasive games. The researchers also noted that the peripherals could change the experience dramatically and this aspect would need separate heuristics.

Song & Lee (2007) studied key factors of heuristics evaluation in games by taking the example of a well-known MMORPG game (World of Warcraft). They conducted both literature review and empirical research for their study and adopted post-surveys and a task oriented analysis. Participants were given specific tasks to follow and usability issues during the play time were noted. The results gathered from these tests were reflected to a new set of heuristics. They based their heuristics on Desurvire's HEP heuristics (Desurvire et al. 2004) and suggested 54 key factors under four key categories; game interface, gameplay, game narrative, and game mechanics.

Desurvire & Wiberg (2009) conducted a research based upon Desurvire's previous approach of using HEP heuristics (Desurvire et al. 2004) and aimed not to just validate but improve it. During the study, HEP heuristics were modified for the game genres such as action, role playing game (RPG), action, adventure and first person shooter (FPS) and discussed with developers working at respected game development companies. After those discussions and refinement of heuristics, PLAY heuristics was proposed. Researchers also set nine categories for general principles of the heuristics: Game Play, Skill Development, Tutorial, Strategy & Challenge, Game/Story Immersion, Coolness, Usability/Game Mechanics and Controller/Keyboard. During the testing, three sets of surveys, depending on the game genre were prepared with a scale of points based on the score of the game that the game received from Metacritic website (www.metacritic.com). Participants were selected from attendees of an annual game conference and were chosen from people who played either the low rated games or high rated ones. In their study, researchers mostly aimed to explain how the PLAY heuristics were defined and how effective they were in a real-world application.

Tan et al. (2010) presented a study conducted to analyze an educational game. In order to achieve this goal, it was aimed to develop a framework of heuristics which is called Instructional Game Evaluation Framework (IGE). The IGE framework had 42 heuristics and was based on “Events of Instruction” method (Gagne & Briggs, 1974), Keller’s ARCS Model for Motivation (Keller, 1999), GameFlow model (Sweetser & Wyeth, 2005) and Nielsen’s heuristics (Nielsen, 1994). 12 primary school students participated in the tests while the research team of five attended as observers and supervisors. An instructional computer game was selected for testing the proposed heuristics. All the students attended pre-explanatory meetings and were given time to play the game without any restrictions. After the play sessions students were divided into three groups and attended focus group meetings. During those meetings, they were asked to comment on the heuristics proposed while the heuristics were simplified for them to understand their notions. The study indicated that including children at the early stages of formal evaluation was effective and valuable since there were revisions coming directly from the participants that effected the heuristics.

Zabion & Shirratuddin (2010) conducted a study focusing on mobile based educational game by proposing a heuristics paradigm with four main modules: Game usability, mobility, gameplay, and learning content. The heuristics were based on Korhonen & Koivisto’s playability heuristics modular approach (Korhonen & Koivisto, 2006) and proposed a module for learning content. In the first phase of the study, participants from primary school students were selected. They were asked to comment on the heuristics and fill surveys. At the second phase a new participant group from 80 exhibition attendees were recruited and asked to play a prototype game. Afterwards, they were asked to fill a Likert scale form representing the heuristics. At the end of the study, researchers revealed results regarding the games performance according to the heuristics they proposed.

Ülger (2013) also proposed a modified version of heuristics, based on Nokia’s Playability Heuristics for Mobile Games. Her study aimed to expand the existing heuristics set for new generation mobile devices and games by adding four different heuristics to the set; distribution of game items, user handedness, use of tilt sensors and

haptic feedback. After the proposal, the heuristics were tested via Game Experience Questionnaire (GEQ) and interviews (Ijsselsteijn et al. 2008). Four mobile games were tested. Two versions of the games were presented during the tests. By inspecting the relevant aspect of the game the heuristic at hand was analyzed. 10 participants were recruited for testing each of the games with a total number of 40.

Khanana & Law (2013) conducted a study to use the Game Flow (Sweetser & Wyeth, 2005) heuristics on digital educational games. They tested four web-based computer games during their study with 100 primary school students. They also re-phrased the heuristics in a way that the students would understand easily so that the heuristics could be given in a survey format. As a result, they indicated the differences among games as well as future possibilities of using heuristics for educational aspects in games, yet they did not propose new heuristics.

Likewise, Rodio & Bastien (2013) conducted a study to evaluate the PLAY heuristics from Desurvire & Wiberg (Desurvire & Wiberg, 2009). In their research 120 amateur e-sport gamers were chosen as participants. A total of three games from different genres (Real Time strategy, MMORPG and FPS) were chosen during the tests. Each participant was asked to fill a five point Likert scale heuristics set which consisted of 47 heuristics. Furthermore, questions regarding the games were directed to the participants in order to receive comments on games they played during the tests. Results indicated that each genre had a different rating of importance in terms of given heuristics, yet the gameplay category carried a generic importance for all of the games.

Another heuristic set evaluation study was carried out by Marciano et al (2014). In their study, they aimed to propose a method for evaluating educational computer games. They used Omar & Jaafar's Playability Heuristic Evaluation for Educational Computer Game (PHEG) (Omar & Jaafar, 2010) in order to analyze an educational computer game. The game developers themselves analyzed the tests and generated a survey format to address heuristics. This study also presented the use of an automated software being employed for the application of heuristic evaluation.

Lastly, Guo & Goh (2016) aimed to use heuristics evaluation instead of anecdotal research on educational games which is a semi-formal method that relies on anecdotal evidence such as user comments. In their study they applied HEP heuristics set to analyze a computer literature game. They presented a survey format of HEP heuristics with 43 items. 39 participants were selected among students. Participants were asked to fill out the survey within a five point Likert scale format after playing the game. Subsequently, the subjective comments of the participants were gathered through interviews. The study proposed suggestions to improve the HEP heuristics set and claimed that two new categories needed to be included; characters and pedagogical effects.

Cross references between heuristics were presented below (Table 2.1) with the notion of listing heuristics which were utilized in more than one research. This limitation of presented heuristics was necessary to avoid listing specific heuristics which involved specific areas of research such as educational context.

**Table 2.2** Cross-Referenced Heuristics of Empirical Evaluation Research on Playability Heuristics

| *Summary of heuristics that are used at least more than one study* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Clear and varied outcomes | x | | | | | | | | | | x |
| Variable difficulty level | x | | | x | | | | | | x | |
| Embodiment of metaphors with physical or other systems that user understands | x | | x | | | | | | | | |
| Audio-visual supports the game | x | | | x | | x | | | | | x |
| Support of a variety of game styles. | x | | | x | x | | | | x | | |
| Using humor appropriately | x | | | x | | | | | | | |
| Making effects of AI visible by ensuring they are consistent with the player's reasonable expectations | | x | | | | | | | | | x |
| Game provides immediate feedback | | x | x | x | x | x | | x | x | | x |
| Player can easily turn the game on/off, and be able to save in different states | | x | | | | | | | | | x |
| The Player experiences the user interface as consistent but the game play is varied. | | x | | x | x | x | | | x | | x |
| Interface/HUD as a part of the game. | | x | | x | | | | | | | x |
| Player has enough information to get started from the beginning | | x | | x | | | | | | | x |
| Context sensitive help | | x | x | x | | | | | | | x |
| Meaningful sounds | | x | | | | | | | | | x |
| Players do not need to use a manual to play game. | | x | | x | | | | | | | x |
| Non-intrusive interface | | x | | | | | | | | | x |
| Make the menu layers well-organized and minimalist to the extent the menu options are intuitive | | x | x | | x | | | | x | | x |
| Quick involvement with tutorials and/or progressive or adjustable difficulty levels | | x | | x | | | | | | x | x |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Art should be recognizable to player, and speak to its function. | | x | | | | | | | | | x |
| Always being able to identify score/status and goal | | x | x | x | x | x | | | x | x | x |
| Standard conventions and natural mapping for controls | | | x | x | | x | | | | | |
| Aesthetic and minimalist design | | | x | | | | | x | | | |
| Clear goals | | | x | x | x | x | | x | x | x | |
| Appropriate rewards for effort and skill development | | | x | x | x | x | | | x | | x |
| Challenge, strategy, and pace are in balance | | | x | x | x | x | | x | x | x | x |
| Fun gaming, without repetitive or boring tasks | | | | x | | x | | | | | |
| Not being penalized repetitively for the same failure | | | | x | | | | | | | x |
| Any fatigue or boredom was minimized by varying activities and pacing | | | | x | | | | | | | x |
| Persistent game world | | | | x | | | | | | | x |
| Application of the newly acquired knowledge / skill | | | | x | x | | | | x | | x |
| Multiple ways to win. | | | | x | | | | | | | x |
| Feeling in control | | | | x | x | x | | x | x | x | x |
| Empathy with the game character | | | x | x | x | | | | x | | x |
| Curiosity and exploration | | x | x | x | x | x | | | x | | x |
| Consistent learning curve with the industry | | | | x | | | | | | | x |
| Screen layout is efficient, integrated, and visually pleasing | | | | x | x | x | | | | | |
| Navigation is consistent, logical and minimalist. | | | | x | x | x | | | x | x | |
| Player error is avoided. | | | | x | x | | | | | | |
| Player interruption is supported | | | | x | | x | | | | | |
| Total concentration | | | | | x | | | x | x | | |
| Storyline relate to your life experiences and grabs interest | | | | | x | | | | x | | x |
| Visuals, animation and music able to capture interest | | | | | x | | | | x | | |
| Font types and sizes used allow easy reading | | | | | x | | | | x | x | |
| Associations of new knowledge and skills with prior knowledge and skills | | | | | x | | | | x | | |
| Learning new concepts and skills | | | x | | x | | | | x | | |
| Feeling of satisfaction and success after gameplay | | | | | x | | | | x | | |
| The game allows to do reflection on learning | | | | | x | | | | x | | |
| The feedback and online help reinforce understanding | | | | | x | | | | x | | |
| Players want to play more of the game | | | | | x | | | | x | | |
| The player does not have to memorize things unnecessarily | | | | | x | x | | | x | | |
| The words and phrases of the game is easy to understand | | | | | x | | | | x | | |
| Feeling confident playing the game | | | | | x | | | | x | | |
| Achieving the learning objectives | | | | | x | x | | | x | | |
| Players easily get help during game play and find this “help” useful | | | x | | x | | | | x | | |
| Warning messages and cues help make less mistakes | | | | | x | | | | x | | |
| The game helps to diagnose players own error | | | | | x | | | | x | | |
| Player can gauge the overall progress at each stage of the game | | | | | x | | | | x | | |
| The game rules assign a final score to the end of each session | | | | | x | | | | x | | |
| The game contains help | | | x | | | x | | | | | |
| The game story supports the gameplay and is meaningful | | | x | x | | x | | | | | x |
| Even if the game cannot be modeless, it should be perceived as modeless | | | x | | | | | | | | x |
| Player experiences fairness of outcomes | | | x | | | | | | | | x |
| | Malone (1982) | Röcker & Haar (2006) | Song & Lee (2007) | Desurvire & Wiberg (2009) | Tan et al. (2010) | Zabion & Shirratuddin (2010) | Ülger (2013) | Khanana & Law (2013) | Rodio & Bastien (2013) | Marciano et al. (2014) | Guo & Goh (2016) |

### 2.2.4 Expert Evaluation

The second approach in the review is conducting expert evaluations using the provided heuristics for evaluating player experience. It was reviewed that a sample group of minimum two participants have performed the evaluations. Even though Nielsen has stated that five experts are normally advised for conducting a heuristic evaluation (Nielsen, 1994), some of the studies have not followed this advice. According to our review, 14 of the relevant articles evaluated either the heuristics or the game via expert evaluation.

Federoff (2002) did a research on existing game heuristics and collated them to analyze the 'fun' aspect of the games. Five people from a game development team were observed and interviewed to suggest a set of heuristics for evaluation of video games. Author analyzed the interviews and observation notes in order to form a list of heuristics. The data collected were compared to formal usability evaluation methods mostly with Nielsen's 10 usability heuristics (Nielsen, 1994). As a result, Federoff presented a set of 10 heuristics for evaluating games yet the suggested heuristics lacked any validation.

Baauw, Bekker & Barengregt (2005) conducted a study on the proposed Structured Expert Evaluation Method (SEEM) which was inspired from Norman's theory-of-action model (Norman, 1988) and Malone's concepts of fun (Malone, 1982). SEEM model was presented to evaluate children's computer games. The aim of the study was mainly to validate the proposed model. They recruited 18 experts from the working area of children, usability and user testing. They also noted that the reason was to improve the SEEM method rather than analyzing the games. Four games were evaluated by the experts for approximately an hour for each game. Experts filled an interaction problem report sheet while conducting the tests. At the end, researchers claimed that the SEEM method was effective in general although they missed several problem categories such as goals, transition and physical action. They also mentioned that SEEM enabled mention some issues which were not revealed from the user-tests done before.

Sweetser & Wyeth (2005), conducted a research on evaluating player enjoyment in video games. They proposed a novel set of heuristics, GameFlow model, for the evaluation of the games utilizing the term 'flow' (Csikszentmihalyi, 1990) at its core. A holistic categorization aimed to evaluate and identify enjoyment in games. They suggested eight key elements including several heuristics in each of them. After suggesting the model, authors validated the model by evaluating two similar real-time strategy games via expert evaluation. As a result, the authors indicated that the model could be used as a guideline for an expert review or basis for other evaluations such as player-testing.

Korhonen & Koivisto (2006) were first to publish playability heuristics for mobile games. They proposed a modular structure for their playability heuristics, which consisted of game usability, gameplay and mobility. Each had distinctive heuristics due to the category and the study was based on literature examination and mobile game reviews. They proposed 29 heuristics in total. Some of the categories and heuristics within those categories were developed from Nokia's Playability Heuristics for Mobile Games. There were two phases of the study. First part involved the use of the three categories of the heuristics with different mobile games. For the first version of the heuristic set, four experts analyzed five mobile games. The experts were either from the field of game design and development or productivity software fields. At the second phase, the set was iteratively improved and the experts conducted the test for the second time, but with different games. According to the results for the study, playability heuristics were effective for evaluating mobile games. Researchers also mentioned that the proposed heuristics could be used in other platforms and games because of its modular structure. Although the heuristics were not compared to previous work and lack empirical validation, playability heuristics and the novel modular structure was well received both in the academia and industry and became basis for other heuristic approaches (Korhonen & Koivisto, 2006). Following their previous work, Korhonen & Koivisto (2007) published a second paper on evaluating mobile multiplayer games. In their latter study they included another module for the multiplayer aspect of mobile games. They prepared the heuristics for the multiplayer category by examining three multiplayer mobile games and literature study.

Köffel & Haller (2008) proposed heuristics for tabletop games. In order to define the heuristics, they incorporated literature reviews and comments from professionals. As a result they presented a modified set of heuristics with 11 items. 12 expert evaluators were asked to evaluate an augmented reality supported tabletop game. There were several sessions during the tests in which the experts were asked to define missing heuristics. In the end, they suggested an iteratively formed heuristic set but mentioned that the last version was not tested. Although this study applied iterative methodology for improving the heuristics, its findings could not be generalized since the focus of the study was on tabletop games.

Korhonen et al. (2009) conducted a study for comparing two playability heuristic sets. In their paper, it was mentioned that the aim was to compare the sets of Korhonen & Koivisto and Desurvire's HEP approach (Korhonen & Koivisto, 2007; Desurvire et al. 2004) since both of those heuristic sets were compatible. During the tests, eight experts were recruited. They were asked to play a mobile game and note the issues about the game in terms of playability. Later they were asked to compare the findings with the given heuristics. The study indicated that playability heuristics had to be improved to be applicable by game developers, in such that the items had to be less in number and more understandable in terms of terminology.

Pinelle et al. (2009) proposed usability heuristics for networked multiplayer games. The study suggested a set of novel heuristics which they called Networked Game Heuristics (NGH). They adapted a previous methodology (Pinelle et al. 2008) which utilized online game reviews to define heuristics. To test the heuristics, 10 experts were asked to play two different games which had multiplayer capabilities via network. The experts were asked to fill out a Nielsen's Severity Scale (Nielsen, 1995). Also the suggested heuristics were compared with the Groupware usability Heuristics [60] during the study. In the result section of their paper, researchers mention that Korhonen & Koivisto's (Korhonen & Koivisto, 2006) playability heuristics were viable in general. Additionally, they mentioned that the aim was to generate a set of heuristics specific to networked games. It was also indicated that the heuristic set was applicable in different platforms and genres, providing a generic property (Pinelle et al. 2009). Other researchers

criticized the article because the previous work which the heuristics were based on (Pinelle et al. 2008) was problematic due to the fact that the online game reviewers were not experts in terms of evaluation or game design. Because of this reason, it is possible to indicate that the suggested set might miss out several aspects of playability.

Koeffel et al. (2010) conducted a study to inspect the use of heuristics to evaluate the overall user experience of games and 'advanced interaction games' (tabletop games). They presented a set of heuristics with three facets; *gameplay*, *game story* and *virtual interface*. The study aimed to develop a set of heuristics which could include more than one aspect of playability and player experience. Researchers based their set of heuristics on Pinelle's (Pinelle et al. 2008) and GameFlow approach (Sweetser & Wyeth, 2005). By conducting an extensive research provided in the literature they put forward 29 items for their heuristics set. The authors claimed that the proposed set included heuristics about only the most important aspects of video games and assumed that it was necessary to investigate the usability/playability of a video game as well as the user experience/player experience to evaluate the overall quality of a game. To determine the effectiveness of the heuristic set, researchers compared the expert evaluation results to common game reviews. Five computer games were tested by two expert evaluators whom were experts in the field of usability and/or games during the tests. Experts were asked to play the games and evaluate them by using the given heuristics set while indicating results via Nielsen's Severity Scale (Nielsen, 1995). The results of the tests (number of issues found through proposed heuristics) were later compared with online review scores. The results indicated that the heuristics were generic though lack the specificity for tabletop games. This study had the authenticity of comparing heuristic evaluation results with common reviews which was referred to Pinelle's approach.

Almeida et al. (2010) conducted a heuristic evaluation of the web-based computer game 'Farmville' by combining heuristics from Federoff (2002), Desurvire (2004) and Pinelle (2008). In their task oriented tests, they indicated 35 heuristics. Each given task during the gameplay was related with certain heuristics. Six evaluators were recruited to attend the tests to fill in the forms with yes or no answers. The study evaluated the game

by only using heuristics and expert evaluations yet the participants lacked the expertise related to the field of gaming or playability or usability.

Suhonen & Vaataja (2010) aimed to study the effect of using modular heuristics on health games. Five previous heuristic sets (Desurvire et al. 2004; Federoff, 2002; Garzotto, 2007; Korhonen & Koivisto, 2006 & 2007) were found to be eligible for being applied during the tests as the authors claimed that these sets complemented each other in terms of given heuristics. After inspection of the heuristic sets, Korhonen & Koivisto's playability heuristics (Korhonen & Koivisto, 2006) were found to be fit for the study. Also, the modular structure of the same study was adopted. Provided that the modular structure was perceived as useful and flexible and could be designed with consideration – given the example that the current heuristics modules could be improved and/or re-arranged. Therefore, to evaluate health games, researchers introduced two new modules to the set, namely for multimodality and persuasiveness. One computer game, one Nintendo Wii game and one mobile game were chosen for the tests. Two experts evaluated the games separately. Experts were asked to fill forms indicating the severity scale and frequency of issues. The results of the study indicated that adding separate modules according to the game genre could be efficient for evaluation purposes as well as being applicable with health games.

Omar & Jaafar (2010) presented the Playability Heuristics for Educational Games (PHEG). They collated the first heuristics set by inspecting the literature for user experience, player experience and pedagogical use in games. Later, experts revised the suggested heuristics and filled a survey for evaluating the PHEG. Experts were also asked to prioritize given heuristic items. As a result of the tests, researchers presented a set with 43 items and five categories with indications of their priority. Researchers indicated that the PHEG set was specifically generated to be used for evaluating educational games hence improving the method by prioritizing the categories. However, their study was not without shortcomings since the heuristics were not examined or verified on an educational game, therefore the study did not involve the empirical validation of the PHEG heuristic set.

Ponnada & Kannan (2012) researched how different mobile games created positive and immersive experiences for the players by using playability heuristics. They based their research on Korhonen & Koivisto's playability heuristic set (Korhonen & Koivisto, 2006). Two expert evaluators were recruited for the examination of each mobile game. Four mobile racing games were chosen for the tests and the experts were asked to play them. After the gameplay, experts filled in the given heuristic forms with yes or no answers. No changes were made from the original heuristics and therefore the study had the value of being a direct implementation of the set. Researchers then compared the results with Android Market ratings and statistics. Researchers indicated that there were positive correlations between heuristic evaluations and statistics from the Android Market only for several games. Because of this reason, they indicated that a more advanced heuristics set had to be developed.

Hynninen (2012) researched the differences between peripherals for first person shooter games using heuristic evaluation. Three games on iPod Touch platform was tested during the study. The author indicated that Pinelle's (2008) heuristic approach was predicated. By reviewing the literature, a new heuristics set was proposed with the focus on first person shooter (FPS) games. Subsequently, the author tested the games using the heuristics to evaluate the iPod Touch games. The result indicated usability issues related to iPod Touch controls.

Carmody (2012) followed a rather different approach. Instead of testing the heuristics, he applied a three session Delphi test process (Skulmoski et al. 2007; Cuhls, 2003) in order to generate collated heuristics. The aim of his study was to investigate which heuristics game designers were considering while developing serious games. Researchers interviewed game designers and generated a first draft for design challenges. After the categorization of the design challenges, they were linked with heuristics cited from the literature. Afterwards, researchers conducted a Delphi survey with 12 expert evaluators and analyzed the first 39 items which were proposed after the first draft. Being iteratively conducted by the tests, 19 items for the heuristic set were suggested as the final result, validated from the literature. In the paper, expert evaluators did not evaluate games

but instead evaluated the proposed heuristics. At the end, the study presented a set of heuristics which could be a guideline for designers developing serious games.

Like Carmody's work, Mohammed & Jaafar (2013), conducted a study on refining the previously explained Playability Heuristics Evaluation for Educational Computer Game (PHEG) heuristic set (Omar & Jaafar, 2010). In the research, 15 expert evaluators were asked to evaluate the PHEG set. The study also aimed to prioritize the categories presented on the PHEG heuristic set. An Analytical Hierarchy Process (AHP) (Saaty, 1990) was applied in order to achieve that goal. At the end, the researchers presented a version of the heuristics set indicating the order of importance for each category.

Wodike et al. (2014) studied the efficiency of empowering teenagers as expert evaluators for analyzing video games in their paper. Based on Pinelle's (2008) heuristic set, they recruited 20 male students as expert evaluators. A mobile game was evaluated during the tests and evaluators were asked to fill in a severity scale form. Even though the study provided results regarding the playability of the game, the researchers highlighted that empowering students as evaluators was non-effective for analyzing the game.

Barbosa et al. (2015) conducted a research about heuristic evaluation of educational games, proposing a blend of items from HEP (Desurvire et al. 2004), PLAY (Desurvire & Wiberg, 2009), and GameFlow (Sweetser & Wyeth, 2005) heuristics. They suggested the set of Heuristic Evaluation for Educational Games (HEEG). The set was applied to five different educational games. Two researchers and one game developer were recruited as expert evaluators during the tests. At the end, the researchers suggested that the set could be a starting point for analyzing specific point of educational games.

Cross references between heuristics were presented below (Table 2.2) with the notion of listing heuristics which were utilized in more than one research. This limitation of presented heuristics was necessary to avoid listing specific heuristics which involved specific areas of research such as educational context.

**Table 2.3** Cross-Referenced Heuristics of Expert Evaluation Research on Playability Heuristics

| *Summary of heuristics that are used at least more than one study* | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Clear and varied outcomes | x | x | | | | | | x | | | | | | | | | |
| Variable difficulty level | x | | x | | | x | | x | x | x | x | | | | x | | |
| Embodiment of metaphors with physical or other systems that user understands | | x | | | | | | | x | | | | | | | | |
| Audio-visual supports the game | | x | | x | | x | | x | x | x | | x | | | | | x |
| Support of a variety of game styles. | | | x | x | | x | | | | | | x | | | | | x |
| Making effects of AI visible by ensuring they are consistent with the player's reasonable expectations | x | | | | | x | | x | | | | | | | | x | |
| Game provides immediate feedback | x | x | x | x | x | x | | | x | x | | x | x | | | | x |
| Player can easily turn the game on/off, and be able to save in different states | x | | x | x | | x | | x | | x | | x | | | | | |
| The Player experiences the user interface as consistent but the game play is varied. | x | | | | | x | | x | | | | | | | | | |
| Interface/HUD as a part of the game. | x | | | | | x | | x | | | | | | | | | x |
| Player has enough information to get started from the beginning | | | | | | x | | x | | | | | | | | | x |
| Context sensitive help | | | | | | x | | | x | x | | | | x | | | x |
| Meaningful sounds | x | | | | | x | | | x | | | | | | | | |
| Players do not need to use a manual to play game. | x | | x | | | x | | | x | x | | | | x | | | x |
| Non-intrusive interface | x | | | | | x | | x | x | x | | | x | | | | |
| Make the menu layers well-organized and minimalist to the extent the menu options are intuitive | x | | | | | x | | x | x | x | | | | | | | x |
| Quick involvement with tutorials and/or progressive or adjustable difficulty levels | x | | x | | | x | x | x | x | x | | | | | | x | x |
| Art should be recognizable to player, and speak to its function. | x | | | | | x | | x | | | | | | x | | | |
| Always being able to identify score/status and goal | | | x | x | | x | | x | x | x | x | x | | x | x | x | x |
| Standard conventions and natural mapping for controls | x | | | x | | x | | x | | x | x | x | | x | x | | x |
| Aesthetic and minimalist design | | | | | | | | | x | | x | | | | x | | |
| Clear goals | | x | x | x | | x | | x | | x | x | x | | | x | | x |
| Appropriate rewards for effort and skill development | x | | x | x | | x | | x | x | x | | x | | | | | |
| Challenge, strategy, and pace are in balance | | x | x | x | | x | | x | | x | x | x | | | x | | x |
| Fun gaming, without repetitive or boring tasks | | | | x | | x | | | | x | | x | | | | | x |
| Persistent game world | x | | x | | | x | | | | x | | | | | | | |
| Multiple ways to win. | x | | x | | | x | | | x | x | | | | x | | | |
| Feeling in control | | x | x | x | | x | | x | x | x | | x | | | | | x |
| Curiosity and exploration | | x | | | | x | | | | | | | | | | | x |
| Screen layout is efficient, integrated, and visually pleasing | | | | x | | x | | | | x | | x | | | | | x |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Navigation is consistent, logical and minimalist. | | | | x | | x | | | | x | x | x | x | | x | | x |
| Player interruption is supported | | | | x | | x | | | | x | | x | | | | | |
| Total concentration | | | x | | | | | | | x | | | | | | | x |
| Storyline relate to your life experiences and grabs interest | | | | | | x | | | | x | | | | x | | | |
| Visuals, animation and music able to capture interest | x | x | x | | | | | | x | x | | | | x | | | x |
| Font types and sizes used allow easy reading | | | | | | | | | | | x | | | | x | | |
| Learning new concepts and skills | x | | | | | | x | | | x | | | | | | | |
| The player does not have to memorize things unnecessarily | | | | x | | x | | | x | x | | x | | | | | |
| Warning messages and cues help make less mistakes | x | | | | | | | | x | x | | | | | | | |
| The game contains help | | | x | x | | x | | | | x | | x | | | | x | |
| The game story supports the gameplay and is meaningful | | | | x | | x | | x | | x | | x | | x | | | |
| Even if the game cannot be modeless, it should be perceived as modeless | x | | | | | x | | | | x | | | | | | | |
| Player experiences fairness of outcomes | x | | | | | x | | x | | | | | | | | | |
| The game has unpredictable yet reasonable story elements. | x | | | | | | | | | x | | | | | | | |
| Multiple goals in each level | x | | | | | | | | x | | | | | | | | |
| Mechanics should feel natural and have correct weight and momentum | x | | | | | | | | | x | | | | | | | |
| Include a lot of interactive props for the player to interact with | x | | | | | | | | | x | | | | x | | | |
| Giving hints (but not too many) | x | | | | | | | | | x | | | | x | | | |
| One reward of playing should be the acquisition of skill | x | | | | | | | | | x | | | | | | | |
| Learning should be fun | | | x | | | | | | | | | | | | | | x |
| Players should become less self-aware and less worried about everyday life or self | | | x | | | | | | | x | | | | | | | |
| Players should experience an altered sense of time | | | x | | | | | | | x | | | | | | | |
| Players should feel emotionally involved in the game | | | x | | | x | | x | | x | | | | | | | |
| Support competition and cooperation between players | | | x | | | | x | | | x | | | | | | | |
| Games should support social interaction between players | | | x | | | | x | | | | | | | x | | | |
| Games should support social communities inside and outside the game | | | x | | | | | | | x | | | | | | | |
| Device UI and game UI are used for their own purposes | | | | x | | x | | | | x | | x | | | | | |
| Visible indicators | | | | x | | x | | x | | x | | x | | | | | |
| Player understands terminology | | | | x | | x | | | | | | x | | | | | |
| Game controls are convenient and flexible | | | | x | | x | | | | | | x | | | | x | x |
| The player cannot make irreversible errors | | | | x | | x | | | | x | | x | | | | | |
| The game accommodates with the surroundings | | | | x | | x | | | | x | | x | | | | | |
| The first-time experience is encouraging | | | | x | | x | | x | | | | x | | | | | |
| Players can express themselves | | | | x | | x | | | | x | | x | | | | | |
| The game does not stagnate | | | | x | | x | | x | | x | | x | | | | | |
| The game is consistent | | | | x | | x | | x | | | | x | | | | | |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| The game uses orthogonal unit differentiation | | | | x | | x | | | | x | | x | | | | | |
| The player does not lose any hard-won possessions | | | | x | | x | | | | | | x | | | | | |
| The cognitive load of the player should not be overburdened | | | | | x | | | | | | | | | | | | x |
| Challenge should be fun | x | | | | x | x | | x | | x | | | | | | | |
| The interpersonal communication and collaboration should be supported by the entirety of the game | | | | | x | | | | | x | | | | x | | | |
| Simple session management | | | | | | | x | | | x | | | | | | | |
| Flexible matchmaking | | | | | | | x | | | x | | | | | | | |
| Appropriate communication tools | | | | | | | x | | | x | | | | | | | |
| Meaningful awareness information | | | | | | | x | | x | | | | | | | | |
| Identifiable avatars | | | | | | | x | x | | | | | | | | | |
| Manage bad behavior | | | | | | | x | | | x | | | | | | | |
| Easy to learn, hard to master | x | | | | | x | | x | | x | | | | | | | |
| The game should be replayable | x | | | | | x | | x | x | x | | | | | | | |
| First action is obvious and gives immediate positive feedback | | | | | | x | | x | | x | | | | | | | |
| The visual representation should allow an unobstructed view of the area that is tied to the location | | | | | | | | x | | | | | | | | x | |
| Allow customization options for controls | | | | | | | | x | x | x | | | x | | | x | x |
| The game should allow customization for different aspects | | | | | | | | x | x | | | | | | | x | |
| Players allowed to build content | x | | | | | | | | | x | | | | | | | |
| The game is paced to apply pressure but not frustrate the player | x | | | | | x | | | | x | | | | | | | |
| Interesting and absorbing tutorial | x | | | | | | | | | x | | | | | | | |
| The design hides the effects of network (in online gaming) | | | | x | | | | | | x | | x | | | | | |
| The game and play sessions can be started quickly | | | | x | | | | | | x | | x | | | | | |
| Maximizes consistency and matches standards | | | | | | | | | | | x | | | | x | | |
| The interactivity of the game is suitable to learners level | | | | | | | | | | | x | | | | x | | |
| The integration of presentation means is well coordinated | | | | | | | | | | | x | | | | x | | |
| The uses of space, color and text are according to the principles of screen design | | | | | | | | | | | x | | | | x | | |
| Quality of user interface is acceptable | | | | | | | | | | | x | | | | x | | |
| Provide specific and self- identified key for specific task (exit, glossary, main, objective) | | | | | | | | | | | x | | | | x | | |
| Overall interface of the game is appealing | | | | | | | | | | | x | | | | x | | |
| The activities are interesting and engaging | | | | | | | | | | | x | | | | x | | |
| The design and the contents are reliable and proven. | | | | | | | | | | | x | | | | x | | |
| Can be used as self- directed learning tools. | | | | | | | | | | | x | | | | x | | |
| Support for self- learning skills. | | | | | | | | | | | x | | | | x | | |
| Medium for learning by doing. | | | | | | | | | | | x | | | | x | | |
| Considers the individual differences. | | | | | | | | | | | x | | | | x | | |
| Performance should be an outcome-based. | | | | | | | | | | | x | | | | x | | |
| Ability to work in their own pace | | | | | | | | | | | x | | | | x | | |

| *Previous Expert Evaluation Research on Playability Heuristics* | Federoff (2002) | Baauw, Bekker & Barengregt (2005) | Sweetser & Wveth (2005) | Korhonen & Koivisto (2006) | Köffel & Haler (2008) | Korhonen, Paavilainen & Saarenpää (2009) | Pinelle, Wong, Stach & Gutwin (2009) | Koeffel et al. (2010) | Almeida, Mealha, Veloso (2010) | Suhonen & Vaataia (2010) | Omar & Jaafar (2010) | Ponnada & Kannan (2012) | Hvnninen (2012) | Carmody (2012) | Mohamed, Jaafar (2013) | Wodike, Sim, Horton (2014) | Barbosa, Rego, Medeiros (2015) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Reliable content with correct flow. | | | | | | | | | | | x | | | | x | | |
| Clear and understandable structure of contents. | | | | | | | | | | | x | | | | x | | |
| Supporting materials are sufficient and relevant | | | | | | | | | x | | x | | | | x | | |
| Materials are interesting and engaging. | | | | | | | | | | | x | | | | x | | |
| Players able to understand the learning goal. | | | | | | | | | | | x | | | | x | | |
| The content is chunk based on topic and subtopic | | | | | | | | | | | x | | | | x | | |
| Major and minor topic is differentiate clearly | | | | | | | | | | | x | | | | x | | |
| Usage of multimedia elements are acceptable | | | | | | | | | | | x | | | | x | | |
| Combination of multimedia elements are adequate | | | | | | | | | | | x | | | | x | | |
| The presentation of multimedia elements are well manage | | | | | | | | | | | x | | | | x | | |
| Suitability of multimedia elements for specific use | | | | | | | | | | | x | | | | x | | |
| Not too many multimedia element in one screen | | | | | | | | | | | x | | | | x | | |
| The use of multimedia elements support meaningfully the text provided. | | | | | | | | | | | x | | | | x | | |
| The quality of multimedia elements (text, image, animation, video and sound) used is acceptable. | | | | | | | | | | | x | | | | x | | |
| The uses of multimedia elements enhance the presentation of information. | | | | | | | | | | | x | | | | x | | |
| Games should provide a lot of stimuli from different sources | | | x | | | | | | | | | | | x | | | x |
| Games should quickly grab the players' attention and maintain their focus throughout the game. | | | x | | | | | | | | | | | | | | x |
| Non-playable content can be skipped | | | | | | | | | x | | | | | | | x | |

### 2.2.5 Inspections

### 2.2.6 Literature Reviews

Another observed approach included articles which solely based on review of existing literature to achieve a more generic point-of-view towards heuristic evaluation.

Schaffer (2007), proposed a white paper for evaluating usability in video games. The aim of the study was to suggest a guideline for evaluating video games via heuristics. It was indicated that with both the utilization of user-tests and expert evaluation methods, it would be possible to analyze the usability of games. With literature review and commendations from the developers, 21 heuristics were suggested with five categories: general, graphical user interface, gameplay, control mapping and level design. Highlighting the lack of empirical research on previous heuristics, the study also did not present test results.

Paavilainen (2010) reviewed video game evaluation heuristics in the context of social games perspective. In the study, a diverse literature review was conducted and four heuristic sets were indicated as comparable among each other (Federoff, 2002; Desurvire et al. 2004; Korhonen & Koivisto, 2006; Pinelle et al. 2009). The focus of the study was social games; therefore a collation of items was prepared from the heuristics mentioned in the study. At the end of the study, the high number of heuristics were criticized and Korhonen & Koivisto's playability heuristics (Korhonen & Koivisto, 2006) was distinguished as the most effective. The author also indicated that user-testing methods combined with heuristics evaluation would provide the most effective analysis. However, the proposed collated set was not tested.

Jerzak & Rebelo (2014), prepared a study for comparing existing heuristics evaluation methods for games with serious games on focus. They also aimed to represent the strengths and weaknesses of existing heuristics in their study. In their paper, they analyzed nine heuristic evaluation approaches. After the elimination of those heuristics, to reach a global view of the related works, authors chose to compare three different

heuristic sets (Federoff, 2002; Desurvire et al. 2004; Desurvire & Wiberg, 2009). They also defined the following three groups/categories for comparison; gameplay, learning & entertainment, usability & game mechanics. The rest of the procedure in the study involved literature inspection and effective aspects for each heuristic set was shown as a result.

### 2.2.7 Game Reviews

Another identified inspection method for developing heuristics is the collection of information from common (online) game reviews which have the potential of offering a much larger sample size.

Livingston et al. (2010), presented a study on using critic reviews of games for refinement of heuristic evaluations. Pinelle's (Pinelle et al. 2008) heuristics were used in the study. Based on previous reviews, authors prioritized the problems which the critics indicated for the games. A modified and genre specific heuristic set was suggested in the study. The authors claimed that by inspecting online reviews, it was possible to prioritize heuristics in terms of severity. Authors also mentioned that even though the study could re-organize the heuristics, it did not encapsulate overall player experience.

Hara & Ovasaka (2014), aimed to develop a heuristic set for action oriented games such as the games developed for *Microsoft Xbox Kinect* peripheral. The study inspected the reviews of 36 motion controlled games with a total number of 256 games. By the inspection of reviews of those games, authors developed new heuristics with 13 items. Although the authors mentioned that there were shortcomings of the use of subjective data gathered from online reviews, there was also the lack of testing the proposed heuristics.

Zhu et al. (2017), utilized the notion of using online reviews to a different level by lexically analyzing 821,122 games with the help of a software. At the end of semi-automated inspections, the authors proposed a set of heuristics and claimed that the studies of Desurvire et al. (2004), Federoff (2002), Malone (1982), Pinelle (2008 & 2009) had deficiencies because of three basic reasons: use of small data sets, depending on

qualitative data and not having been empirically testing, and lastly focusing on small number of games and therefore not being generic. As a result the authors presented 90 playability heuristics.

Cross references between heuristics were presented below (Table 2.3) with the notion of listing heuristics which were utilized in more than one research. This limitation of presented heuristics was necessary to avoid listing specific heuristics which involved specific areas of research such as educational context.

**Table 2.4** Cross-Referenced Heuristics of Research via Inspections on Playability Heuristics

| *Summary of heuristics that are used at least more than one study* | | | | | | |
|---|---|---|---|---|---|---|
| Support of a variety of game styles. | | | | | x | x |
| Making effects of AI visible by ensuring they are consistent with the player's reasonable expectations | | | x | x | | x |
| Game provides immediate feedback | | | x | | x | x |
| Context sensitive help | | | x | | | x |
| Meaningful sounds | | | x | | | x |
| Non-intrusive interface | | | x | | | x |
| Quick involvement with tutorials and/or progressive or adjustable difficulty levels | | x | x | x | | |
| Always being able to identify score/status and goal | x | | x | x | | |
| Standard conventions and natural mapping for controls | x | | x | | x | x |
| Clear goals | x | | x | | | |
| Appropriate rewards for effort and skill development | | x | x | | | |
| Challenge, strategy, and pace are in balance | | | x | | | x |
| Fun gaming, without repetitive or boring tasks | x | | x | | x | x |
| Persistent game world | | x | x | | | |
| Feeling in control | x | | x | | | |
| Storyline relate to your life experiences and grabs interest | | | x | | | x |
| The game story supports the gameplay and is meaningful | | | x | | | x |
| Visible indicators | x | | | | | x |
| Game controls are convenient and flexible | | | x | x | | x |
| The player cannot make irreversible errors | x | | x | | | |
| The game should be replayable | | | x | | | x |
| The game should allow customization for different aspects | | | | x | | x |
| *Previous Research on Playability Heuristics via Inspections* | Schaffer (2007) | Paavilainen (2010) | Jerzak & Rebelo (2014) | Livingston, et al. (2010) | Hara & Ovasaka (2014) | Zhu et al. (2017) |

### 2.2.8 Mixed Method

The last observed method involved studies with mixed-method modality, combining empirical research, expert evaluation and/or inspection methods.

Desurvire et al. (2004), proposed the Heuristics of Playability (HEP) framework and prepared a heuristics set of 43 items, based on literature and reviewed by several experts. The expert evaluator formed the HEP set while focusing on how each heuristic was indicating a playability issue. The HEP heuristics set consisted of four categories; gameplay, game story, mechanics and usability. This model was tested via a prototype game. During the study, the researchers conducted user-testing method for validating and comparing the results from the heuristic evaluation. The heuristics were analyzed and evaluated through four participants in two-hour long sessions where they played the game and evaluated the aspects of the game using the given heuristic items. The evaluator logged the actions and observation notes during the testing period. The user-tests included think-aloud play sessions and satisfaction questionnaires as well as observation notes taken by the supervisors. At the end of the study, both the results from user-tests and HEP evaluations were compared and the overall findings indicated that HEP heuristics were much more effective for finding issues related to the playability of the game rather than user-tests. This study proposed a new set of heuristics which was then used by several researchers but it had problematic aspects such as the unclear wordings for heuristics statements.

Pinelle et al. (2008), presented a study on evaluating early versions of the games via heuristics. They utilized 108 online game reviews to form 10 heuristics. Subsequently, 10 more heuristics were added to involve multiplayer aspects of the games. Proposed heuristics were prepared with reference to released game reviews. After gathering the reviews, authors presented 12 problem categories. Finally, preliminary tests were conducted with the suggested set of heuristics. Five participants were asked to test an under-developed computer game using the heuristics. The participants were asked to fill a report form and Nielsen’s Severity Scale. Authors emphasized that the set could identify usability issues of the game and the study offered a novel approach for using online game

reviews as basis for defining heuristics. Even though the heuristic set had insufficient preliminary evaluations, the study received criticism from other researchers because of the use of biased online reviews.

Febretti & Garzotto (2009), conducted a research on long-term engagement effects of games and the relation of usability. The aim of their study was to determine the effects of the game interface and its relation to the long-term player experience. To achieve that goal, they based their approach on the comparison of usability and playability. By using inspection method, they blended and modified 22 heuristic items from the literature and presented the set in seven categories. Likewise, they followed a similar method for generating a usability focused heuristic set with five categories and 14 items. They applied both user-testing and expert evaluation methods in their study. They tested eight commercial long term games with eight groups of participants with a total number of 47 participants. The tests also involved 20 game design experts and inspectors of usability and playability. To evaluate the aspect of engagement, they conducted user-tests with and without supervisors. Ultimately, they investigated the correlations from both test results and claimed the study had the intrinsic value of focusing on engagement. They indicated that playability heuristics had higher correlation values rather than usability heuristics. The result of the study offered a methodological approach in general.

Papaloukas et al. (2009), conducted a study with a multi-modal methodology, combining user-tests with expert evaluations. Since there were no adequate methods or methodologies for evaluating a game's usability, they proposed using a modified set of heuristics based on Nielsen's heuristics (1994). They conducted tests on 2 different games in different platforms (*Nintendo Wii* & web-based computer game). For the user-tests, 30 participants were selected for usability evaluations. A specialized software was used for gathering metrics including user logs, facial expressions and verbal reactions using a camera. Player actions were recorded and analyzed by three usability experts. On the other hand, experts played the games for a week and wrote down the heuristics they used to identify the problem. Authors resulted their study by indicating the importance of the combination of these two methods, noting that the final results were enriched with the data gathered during player observations.

Jegers (2009), studied on defining the enjoyment in pervasive games. Three pervasive tabletop games were tested using the *GameFlow* model (2005) in three phases. The first phase of the research involved user-testing with 58 participants. The second phase involved six expert evaluators testing the heuristics. Lastly, the third phase involved sessions with both groups conducting a playtest and a focus group study. The author presented 14 new heuristics to be added to *GameFlow* model.

Desurvire & Wixon (2013), aimed to determine the effectiveness and advantages of using heuristics for evaluating video games in their study. The focus on the study was to identify differences between the findings provided by heuristics and informal usability inspections. In their research, both the *PLAY* (2009) and Game Approachability Principles (*GAP*) heuristics (2015) sets were analyzed. Two browser based computer games were evaluated by 22 experts from the fields of game development and game review in three sessions each. At first, evaluators were asked to perform informal evaluations without heuristics, later with using *PLAY* heuristics and lastly *GAP* heuristics. Experts were asked to mark their comments by coded representations. The overall results indicated that utilization of heuristics during the evaluations help not only spot problems and suggest solutions but also help participants recognize effective elements of the design and suggest improvements. The researchers suggested that both sets were not only sufficient for analyzing the games but also effective for generating suggestions related to the issues in the gameplay. The mean frequency of issues mentioned during the tests were higher for heuristic evaluation compared to informal evaluations. It was noted that using heuristics provided more issues and thus was a better choice for evaluation than previously conducted informal evaluations.

Desurvire & Wiberg (2015), aimed to compare different evaluation methods to test *GAP* heuristic set. Also, they aimed to test this new set of heuristics on different gaming platforms. Researchers utilized usability and heuristics evaluation techniques to compare them. One researcher applied heuristics evaluation method utilizing heuristics gathered from playability and usability literature while the other applied user-tests. Four games were tested during the study. After the tests, researchers analyzed and compared the results

from both methods. 32 participants attended to the empirical tests. In the result section, researchers claimed that *GAP* heuristics and user-tests supported each other while indicating the best approach for analyzing the overall experience in games was the use of both methods simultaneously. Desurvire noted that, like *PLAY* heuristics, *GAP* principles held a guiding purpose therefore not directly aimed to evaluate playability.

Hochleitner et al. (2015), introduced a heuristic framework for evaluating user experience in games. The study aimed to improve previously presented heuristic approaches and correlated them with common game reviews. The study was complementary to Koeffel's research (2010). In order to measure the applicability of the heuristic framework, six games were tested. The online game review ratings were later compared with the results of heuristic evaluation. The proposition of the heuristics was based on the previous works of Malone (1982), Federoff (2002), Desurvire (2004), Shaffer (2007), Pinelle (2008), Koeffel (2010), and Korhonen & Koivisto (2006). However, the focal point of the suggested new heuristic set was Koeffel's (2010) set with 29 items. At the end, a total set of 49 items was proposed. The games tested were selected due to online game review ratings and evaluated by three expert evaluators who had previous experience with the heuristics set. Consequently, it was stated that there was a correlation between average game review ratings and results obtained from the heuristics study.

Cross references between heuristics were presented below (Table 2.4) with the notion of listing heuristics which were utilized in more than one research. This limitation of presented heuristics was necessary to avoid listing specific heuristics which involved specific areas of research such as pervasive games context and provide a holistic point of view.

**Table 2.5** Cross-Referenced Heuristics of Mixed Method Research on Playability Heuristics

| *Summary of heuristics that are used at least more than one study* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable difficulty level | x | | | | x | x | x | x |
| Audio-visual supports the game | | | | | | x | x | x |
| Support of a variety of game styles. | | | x | | x | x | x | |
| Using humor appropriately | | | | | | x | x | |
| Making effects of AI visible by ensuring they are consistent with the player's reasonable expectations | x | x | x | | | | | x |
| Game provides immediate feedback | x | | | | x | x | x | x |
| Player can easily turn the game on/off, and be able to save in different states | x | | | | x | | | x |
| The Player experiences the user interface as consistent but the game play is varied. | x | | | | | x | x | |
| Interface/HUD as a part of the game. | x | | x | | | x | x | x |
| Player has enough information to get started from the beginning | x | | | | | x | x | |
| Context sensitive help | x | | | | | x | x | x |
| Players do not need to use a manual to play game. | x | | | | x | x | x | |
| Non-intrusive interface | x | | x | | | | | |
| Make the menu layers well-organized and minimalist to the extent the menu options are intuitive | x | | | | | | | x |
| Quick involvement with tutorials and/or progressive or adjustable difficulty levels | x | x | | | x | x | x | x |
| Always being able to identify score/status and goal | x | x | x | | x | x | x | |
| Standard conventions and natural mapping for controls | | | | x | | x | x | x |
| Clear goals | x | | x | x | x | x | x | x |
| Appropriate rewards for effort and skill development | x | | x | | x | x | x | x |
| Challenge, strategy, and pace are in balance | | | | x | x | x | x | x |
| Fun gaming, without repetitive or boring tasks | | | | | | x | x | x |
| Persistent game world | x | | | | x | x | x | x |
| Multiple ways to win. | x | | | | x | x | x | |
| Feeling in control | x | | | | x | x | x | x |
| Application of the newly acquired knowledge / skill | x | | | | | x | x | |
| Visuals, animation and music able to capture interest | | | | x | | x | x | |
| Warning messages and cues help make less mistakes | | x | | | x | | | |
| The game contains help | x | | | | | | | x |
| The game story supports the gameplay and is meaningful | x | | | | | x | x | x |
| Multiple goals in each level | x | | | | x | | | |
| Players should feel emotionally involved in the game | | | | | x | | | x |
| The cognitive load of the player should not be overburdened | x | | | x | | | | |
| Meaningful awareness information | x | | | | | x | x | |
| Easy to learn, hard to master | | | | | | x | x | x |
| The players should not lose any hard won possessions. | | | | | | x | x | |
| The game should be repayable | x | | | | | | | x |
| First action is obvious and gives immediate positive feedback | | x | x | x | | | | x |
| The game is paced to apply pressure but not frustrate the player | x | | | | | | | x |
| Curiosity and exploration | | | | | | x | x | |

| *Previous Mixed Method Evaluation Research on Playability Heuristics* | Desurvire, Caplan, Toth (2004) | Pinelle, Wong, Stach (2008) | Febretti & Garzotto (2009) | Papaloukas et al. (2009) | Jegers (2009) | Desurvire & Wixon (2013) | Desurvire & Wiberg (2015) | Hochleitner et al. (2015) |
|---|---|---|---|---|---|---|---|---|
| If there is a game story, the player is eager to spend time thinking of the possible outcomes. | x | | | | | x | x | |
| Not being penalized repetitively for the same failure | | | | | | x | x | |
| Game control should allow a smooth gaming experience without unnecessary pauses | x | | | | | x | x | |
| Provide consistency between the game elements and the overarching setting and story to suspend disbelief. | x | | | | | | | x |
| Empathy with the game character | x | | | | | x | x | |
| The game offers something different in terms of attracting and retaining the players' interest. | x | | x | | | x | x | |
| Player error is avoided | | | | | | x | x | |
| Consistent learning curve with the industry | | x | x | | | x | x | |
| Games should respond to users' actions in a predictable manner | | x | | | | | | x |
| The game should provide views that allow the user to have a clear, unobstructed view of the area | | x | | | | | | x |
| Allow users to skip non-playable and frequently repeated content | | x | | | | | | x |
| Allow customization options for controls | | x | | x | | | | |
| Game controls are convenient and flexible | | | x | | x | | | |
| The game should provide different challenge levels for different players | | | x | x | x | | | x |
| Player interruption is supported | | | | | | x | x | |
| Skills are useful | | | | | | x | x | |

### 2.2.9 Implications

By categorization of methodological differences between heuristic evaluation researches in this study, it was aimed to present a novel perspective to the domain of playability evaluations. It was also aimed to offer a holistic view to provide a guide for future research regarding methodological approaches for heuristic evaluation of games. The researches indicated that most of the authors suggested using more than one method in order to validate the proposed heuristics. It was observed that literature review for defining the playability heuristics was the most common way to conduct studies.

This chapter represented that 15 heuristics were common for all of the methods (Table 2.5). Accordingly, the review indicated that studies combining heuristic sets were efficient but lacked validation. Multi-modality in research, such as using user testing in order to validate expert evaluations, yielded more results in comparison with other methods. Further research might involve using collated playability heuristics via expert evaluations and empirical evaluations in order to validate the provided heuristics.

**Table 2.6** Common Playability Heuristics in the Literature

| Heuristics | Number of References |
|---|---|
| Support of a variety of game styles. | 15 |
| Making effects of AI visible by ensuring they are consistent with the player's reasonable expectations | 13 |
| Game provides immediate feedback | 27 |
| Context sensitive help | 14 |
| Non-intrusive interface | 12 |
| Quick involvement with tutorials and/or progressive or adjustable difficulty levels | 22 |
| Always being able to identify score/status and goal | 29 |
| Standard conventions and natural mapping for controls | 21 |
| Clear goals | 26 |
| Appropriate rewards for effort and skill development | 21 |
| Challenge, strategy, and pace are in balance | 25 |
| Fun gaming, without repetitive or boring tasks | 14 |
| Persistent game world | 13 |
| Feeling in control | 23 |
| The game story supports the gameplay and is meaningful | 16 |

# 3. EXPLORING THE PLAYER EXPERIENCE IN DIFFERENT GAMING PLATFORMS

Mayra (2008) mentions that every game necessitates a core (gameplay) and a shell (symbolic presentations) that operate on a gaming platform. During play, players interact with game mechanics through a user interface providing immersive interactivity for the player (Schell, 2008). This interface is considered to be consisted of input devices to control the game and/or the avatar, and output devices that provide stimuli about the game to the player (Clanton, 2000). When inspected from a holistic perspective, the most crucial aspect of the games, the interaction, is clearly dependent on the platform and its features. Korhonen describes gaming platforms as; "*gaming platform consist of the available processing power, memory, a graphical accelerator, network connection or other features of the platform*" (Korhonen 2016). Although this definition is valid to some extent, for most of the games the gaming platform is the determinant factor for the developers while designing games. Essentially, the gaming platform enables interaction between the player and the gameplay. Hence it is crucial to understand the role of platforms in player experiences.

## 3.1 Domain Specific Playability Heuristics

As discussed and identified in the previous chapter, studies conducted for evaluating player experience via heuristics were diverse and fundamentally different between each other. Either the method of application or the goals of the heuristics vary. Similar to these variations, several researches among previous literature have focused on developing and applying platform or game genre specific heuristics for evaluating player experience.

Although he was researching the elements for an entertaining interface design, Malone (1982) was the first to conduct a study on proposing heuristics to be used in learning and teaching. He presented a set of heuristics for instructional games and suggested that there were three main heuristics for achieving entertaining interfaces. Following Malone's work, Clanton (1998) proposed 15 design principles that are considered important for game design. However they were not developed in the form of 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 collected these principles via checking various genres of games including strategy, adventure, puzzle, role-playing, and action genres. The inclusive tests were fundamentally important to propose a generalizable set of principles.

In terms of game types and genres, Röcker & Haar (2006) investigated the use of existing heuristics on evaluating pervasive games. In their study, they indicated that there is a need for improvement to the existing heuristic set for analyzing pervasive games and noted that game peripherals needed separate heuristics. Korhonen & Koivisto (2006) were

the first to propose playability heuristics focusing on mobile platform. They proposed their modular approach based on existing playability heuristic research. They proposed a modular structure for utilizing the mobility aspect of mobile games. Pinelle et al. (2008) have proposed a novel usability heuristics set for networked multiplayer games which they called *Networked Game Heuristics*. Adapting game reviews as a baseline, they also indicated that Korhonen & Koivisto's (2006) playability heuristics were viable to some extent for analyzing games yet it was not sufficient for analyzing networked games. Köffel & Haller (2008) have also presented a set of heuristics but specifically for tabletop games and not applicable to other game genres. Similar to Röcker & Haar (2006), Jegers (2009) also utilized existing heuristics for evaluating pervasive games and proposed 14 novel heuristics to be added on the existing *GameFlow* (Sweetser & Wyeth, 2005) heuristic set. Tan et al. (2010) presented a framework for analyzing educational games. In their work, they proposed the *Instructional Game Evaluation Framework* with 42 heuristics based on existing heuristic sets. Zabion & Shirratuddin (2010) have studied focusing on mobile based educational game, utilizing playability heuristics proposed by Korhonen and Koivisto (2006). After their study, they proposed a new module for the existing set, learning content. Suhonen & Vaataja (2010) have conducted a study on health games at the same year with the emphasis on using modular structure for heuristics. After inspecting existing heuristic sets they identified the playability heuristics (Korhonen & Koivisto, 2006) to be most eligible. They indicated that adding separate modules according to the game genre could be efficient for evaluation purposes. Omar & Jaafar (2010) have presented a study on educational games. They presented the *Playability Heuristics for Educational Games* based on previous research. It was indicated that the set was specifically designed for analyzing educational games. Paavilainen (2010) compared heuristic sets while social games were in focus. A collated set of heuristics was proposed while indicating that playability heuristics (Korhonen & Koivisto, 2006) was most effective. Later, Mohammed & Jaafar (2013) have presented an improved version for the previously presented set. Hynninen (2012) have evaluated differences between different peripherals for first person shooter games. A novel set of heuristics was proposed focusing on first person shooter game genre. Jerzak & Rebelo (2014) have conducted a research on evaluating serious games, blending various existing heuristics. Hara & Ovasaka (2014) have proposed a novel heuristic set focused on action oriented games (for ex. Kinect

games developed for Microsoft Xbox console platform). Barbosa, Rego & Medeiros (2015) have also collated several existing heuristic sets for evaluating educational games. Guo & Goh (2016) have conducted a study for proposing heuristics for educational games. They chose to employ an existing heuristic set from Desurvire (2004) and improve it by implementing additional heuristics.

Table 3 presents a summary of heuristics sets that have been published focusing on domain-specific analysis. Even though the gaming platform hold a crucial role regarding the player experience as indicated in the begging of the section, none of the proposed heuristics refer to platform differences. It can be observed that mostly the domain of educational games was targeted. Additionally it can be seen that the improvement of heuristic sets are yet to cease since these studies indicate the emerging need for novel heuristics regarding specific domains.

**Table 3.1** Domain-specific Playability Heuristic Studies

| Author (s) | Year | Domain |
|---|---|---|
| *Malone* | 1982 | Educational Games |
| *Röcker & Haar* | 2006 | Pervasive Games |
| *Korhonen & Koivisto* | 2006 | Mobile Games |
| *Pinelle et al.* | 2008 | Networked Multiplayer Games |
| *Köffel & Haller* | 2008 | Tabletop Games |
| *Jegers* | 2009 | Pervasive Games |
| *Tan et al.* | 2010 | Educational Games |
| *Zabion & Shirratuddin* | 2010 | Mobile Educational Game |
| *Suhonen & Vaataja* | 2010 | Health Games |
| *Omar & Jaafar* | 2010 | Educational Games |
| *Paavilainen* | 2010 | Social Games |
| *Hynninen* | 2012 | Game Peripherals |
| *Mohammed & Jaafar* | 2013 | Educational Games |
| *Jerzak & Rebelo* | 2014 | Serious Games |
| *Hara & Ovasaka* | 2014 | Action Oriented Games |
| *Barbosa, Rego & Medeiros* | 2015 | Educational Games |
| *Guo & Goh* | 2016 | Educational Games |

### 3.2 Analyzing Platform Differences

Game researchers have questioned the applicability of heuristic evaluation methods for analyzing playability on video games. Moreover, most of the proposed heuristic

evaluation approaches are not validated empirically. The potential of evaluating playability of games necessitate improvements for analyzing specific features of different gaming platforms.

In terms of analyzing player experiences between platforms, Suhonen & Vaataja (2010) have conducted a study based on modular playability heuristics work in the evaluation of different types of health games. They compiled five sets of heuristics that they consider complimentary between each other (Desurvire et al. 2004; Federoff, 2002; Garzotto, 2007; Korhonen & Koivisto, 2006, 2007) and developed a new set of heuristics for evaluating health games. Accordingly they defined six modules for their heuristics set: *1) game usability, 2) gameplay, 3) game story, 4) content and education, 5) social interaction and multi-playing*, and *6) mobility*. They conducted the evaluations on three different multiplayer exercise games on PC, mobile and Nintendo Wii platforms. Researchers noted that these games were selected according to their similarity on gameplay and mechanics and the differences between their gaming platforms. A different set of heuristics was employed for each game and was played by two usability researchers. After the tests, researchers perceived that the modular structure of heuristics was useful and efficient. More importantly, researchers mentioned that despite the differences in the tested platforms of games and games themselves, the suitability of the heuristics for evaluating each of them received relatively similar comments. This indicates that the heuristic evaluation by itself was found to be not sufficient for evaluating platform differences but was suitable for analyzing the games. The lack of heuristics for identifying platform differences was highlighted by the researchers themselves. They indicated that a dominant element of Wii platform, the use of various and rather different controllers were not effectively evaluated and was clear "*that without any heuristics on multimodality this aspect of the game was almost completely left out of the evaluation, even though some pieces relating to the multimodal gaming style are present in the heuristics on game usability and gameplay*". Afterwards, they suggested that a novel set of heuristics covering multimodality of the player experience should be added.

Interested in comparing player experience between controllers, Gerling et al. (2011), have conducted a study on measuring the impact of game controllers on player experience

by focusing on first person shooter games. They emphasized the importance of multi-platform for games and conducted a study with 45 participants to analyze player experience and playability issues between platforms. As distinct from previous studies, they explored the effects of players being forced to switch to an unfamiliar platform. In their study, they compared PC and console platforms using a first person shooter game, *Battlefield: Bad Company 2*. To conduct the tests, the researchers have employed the Game Experience Questionnaire (GEQ) (Ijsselsteijn et al. 2008). Furthermore they administered the Game Engagement Questionnaire (GEnQ) (Brockmyer et al. 2009) which was designed to explore player engagement in video games. According to the results of the questionnaires, the researchers concluded that the players switching to a new platform, experience more usability issues and consider themselves more challenged, but report an equally positive overall experience as players on their comfort platform. In this context, researchers have suggested that, rather than focusing on adjusting first person shooter games particular to a gaming platform, designers could focus more on making generally enjoyable game mechanics. Although this might be the case for the tests, researchers noted that additional exploration of the impact of different gaming platforms should be done.

Focusing on game controllers, Zaman et al (2010) have conducted a similar research by comparing touchscreens with traditional controllers in the context of handheld gaming. They presented a study which compares touchscreen-based controls and physical controls using Ubisoft's *Assassin's Creed: Altair's Chronicles*. They utilized two hand-held platforms during the tests comparing virtual and physical buttons while 12 participants completed a game level four times in each platform. Researchers have recorded the sessions and took notes of the game metrics during the play sessions. They conducted task oriented tests which involved completing a level of the game on a given platform. As a result, the researchers indicated that haptic feedback afforded by the physical controllers resulted in better control over touchscreen controls. To conclude they suggested that the games should be designed according to the given platform capabilities and not emulate controls that would simply do not work on touchscreens in this example.

Raffaele et al. (2015) have also conducted a study on evaluating multi-platform games. They evaluated a casual game, "Doctor Who: Legacy", which is available for PC and mobile platforms. They noted that, casual games were important because of their similar achievement for good results compared to consoles and computers. In addition authors mention that, unlike console or computer games, casual games do not require commitment in terms of time and dedication as well as skillful players. In order to evaluate the game, they used the approach from Preece et al. (2005) and utilized the concepts of; being fun, easy to understand, efficient to use, aesthetically pleasing and motivating. Accordingly, they compared two platforms. To conduct a survey, they recruited 10 people at random, aged between 14 and 28 years old. Players were asked to play the game in both platforms on mobile and PC respectively. Players were also observed during play sessions. At the end of the tests, players were asked to answer a questionnaire with 14 questions. To conclude, researchers indicated that the tests did not find differences in terms of visual and physical interactions and note that future research should focus more than the properties of interface when comparing platforms.

Kokil & Sanchez (2015) investigated the application of *Facets of Playability* approach for analyzing differences between PC and tablet gaming. They based their evaluations on several evaluation approaches including *Self-Assessment Manikin, PIFF2,* Game Experience Questionnaire (GEQ) (Ijsselsteijn et al. 2007) and *Facets of Playability* to measure the game enjoyment during their study. By triangulating physiological data of emotional responses using galvanic skin response, heart rate, and subjective feelings data of facets of player experience, they aimed to identify the effects of different platforms in terms of controls. A total of 14 players were recruited for the tests. A demographic survey and a mood questionnaire was used before the play sessions to gain knowledge about the players. Although they conducted the test, only the preliminary results were published in their study indicating that mode of interaction (i.e. touch screen) brings a different kind of user experience that is not achieved by other platforms.

Although there are several studies on evaluating player experience, there are only a number of emerging studies that focus on platform differences in the context of player experience. Previous studies have shown that employing an evaluation method

individually was not efficient and resulted in lack of holistic analysis for evaluating platform differences. Additionally, employing task-oriented tests and utilizing metrics and/or logs of the players provided success rating and did not provide sufficient results when analyzing the experience of the players.

### 3.3 Conducting the Multi-modal Evaluation

In this thesis, I aim to analyze and inspect playability evaluation methods and describe methods to be used for evaluating player experiences between different platforms. To assess the platform differences it is obvious that individual evaluations were not efficient and holistic to gain an inclusive understanding of the player experience. Therefore player evaluations were performed during the research including pre-test questionnaires, experience surveys, post-tests, interviews and observations. As described in the previous chapter, a mixed-method approach was adapted during the thesis research for each study. Accordingly, three different methodological approaches were employed for analyzing player experiences have been conducted during this research. These approaches were decided to be administered during the research because of their widely-used and inclusive properties.

In this section, we will take a look at these approaches that have been utilized during the research to help evaluate the playability of games in different platforms.

#### 3.3.1 Playability Model Framework

Sanchez et al. (2009a) propose a rather holistic definition and framework for playability where the focus is on the user and player experience instead of the game. They analyzed and categorized features of games and proposed the *Playability Model (PM)*. PM framework has the distinction of collating various heuristics into the framework of playability. This model aimed to characterize playability of games and included seven attributes: *Effectiveness, Learnability, Immersion, Satisfaction, Motivation, Emotion* and *Socialization*. Researchers have characterized the playability by utilizing existing knowledge of usability and gave different meanings in the context of games (Figure 3).

The model framework is inspired from the three heuristics of usability; e*ffectiveness, efficiency* and *satisfaction* as well as their correlations with playability.

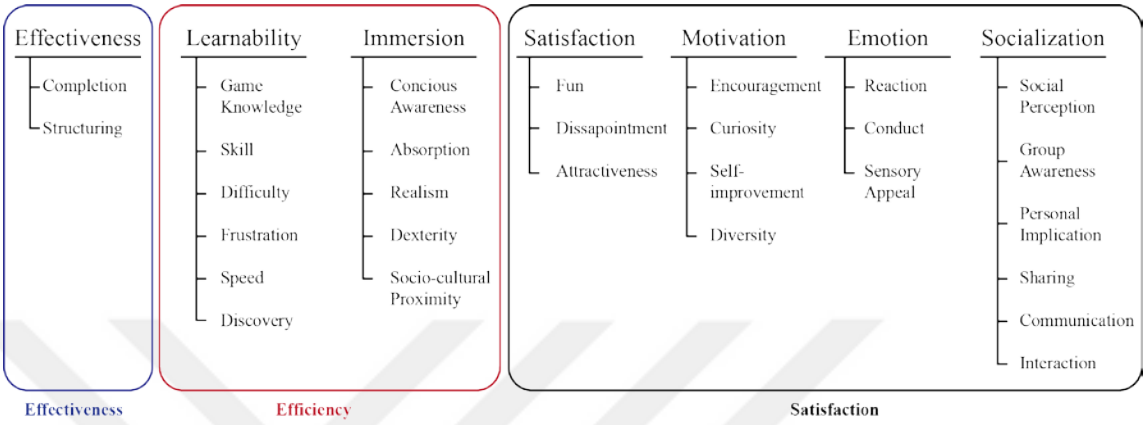



**Figure 3.1** Playability Model: Attributes and Properties

***Satisfaction:*** The researchers defined this as "*the gratification or pleasure derived from playing a complete video game or from some aspect of it.*" They characterized it as using the following categories: *Game Fun*, *Contents*, *Disappointment* and *Attractiveness*.

***Learnability:*** Researchers defined this as *"the player's capacity to understand and master the game's system and mechanics"*. They characterized it with: *Game Knowledge, Player Skill, Game Difficulty, Player Frustration, Speed of Learning* and *Discovery Techniques.*

***Effectiveness****:* The researchers defined this item as "*the time and resources necessary to offer players a fun and entertaining experience whilst they achieved the game's various objectives and reach the final goal*". They identified it with the properties of: *Game Completion* and the *Structuring of the Game Resources.*

***Immersion:*** The researchers have identified this item as *"the capacity of the video game contents to be believable, such that player becomes directly involved in the virtual game world"*. They characterized the item with the properties of: *Conscious awareness, Absorption in game, Game Realism, Control Dexterity* and *Socio-Cultural Proximity with the game*.

***Motivation:*** The researchers defined this as "*the set of game characteristics that prompt a player to realize specific actions and continue undertaking them until they are*

*completed*". They characterized the item with: *Encouragement Techniques, Curiosity about the game, Player Self-Improvement* and *Diversity of game resources.*

***Emotion:*** They refer to this item to the *"player's involuntary impulse in response to the stimulus of the video game that induces feelings or a chain reaction of automatic behaviors"*. They characterized this item with the properties of: *Player Reaction, Game Conduct* and *Sensory Appeal for game elements.*

***Socialization:*** Researchers defined this as *"the set of game attributes, elements and resources that promote the social dimension of the game experience in a group scenario"*. They proposed the following properties defining the item: *Social Perception, Group Awareness, Personal Implication, the Sharing of the Social Resources, Communication Techniques* and *Interaction Rules of the socialization game.*

Moreover, as to illustrate how to evaluate and analyze the player experience using the PM framework, Sanchez et al. (2012) conducted a research on analyzing experience in video games. During the study, researchers employed a mixed-method approach, utilizing playability model heuristic evaluations and user tests with questionnaires and observations. At the end of their study, they indicated that the proposed playability attributes were useful for describing the experience from the UX point of view.

### 3.3.2 Game Experience Questionnaire

Even though there is a lack of operationalizing and measuring the validity and reliability of player experience evaluation methods, a reliable instrument was proposed during the "*Fun of Gaming*" (FUGA) project (Poels, de Kort & Ijsselsteijn, 2007). The proposed Game Experience Questionnaire (GEQ) is a self-report measure which was composed of three questionnaire modules: C*ore Module, Social Presence Module* and *Post-Game Module*. The *Core Module*; is defined as the heart of the questionnaire, probing several components of the game experience. The *Social Presence Module*; is developed to probe player experience and involvement with other players. Lastly, the *Post-game Module*; is defined as the questionnaire aimed to inspect players experience after game sessions. In this framework, components such as *immersion, tension,*

*competence, flow, negative affect, positive affect,* and *challenge* are defined as important factors to offer an understanding and evaluate player experience in games.

The Core Module has 33 items and 7 different components: *Competence* (e.g. 'I felt skillful', 'I felt competent'), *Sensory and Imaginative Immersion* (e.g. 'I was interested in the game's story', 'I felt imaginative'), *Flow* (e.g. 'I was fully occupied with the game', 'I lost track of time'), *Tension/Annoyance* (e.g. 'I felt annoyed', 'I felt irritable'), *Challenge* (e.g. 'I thought it was hard', 'I felt pressured'), *Negative Affect* (e.g. 'It gave me a bad mood', 'I found it tiresome') and *Positive Affect* (e.g. 'I felt content', 'I felt happy').

The Social Presence Module has 18 items and 3 different components: *Psychological Involvement – Empathy* (e.g. 'I empathized with the other(s)', 'I felt connected to the other(s)'), *Psychological Involvement – Negative Feelings* (e.g. 'I felt jealous about the other(s)', 'I influenced the mood of the other(s)'), *Behavioral Involvement* (e.g. 'My actions depended on the other(s) actions',' The other's actions were dependent on my actions').

Lastly, the Post-game Module has 17 items and 4 different components: *Positive Experience* (e.g. 'I felt revived', 'It felt like a victory'), *Negative Experience* (e.g. 'I felt bad', 'I felt guilty'), *Tiredness* ('I felt exhausted', 'I felt weary'), *Returning to Reality* (e.g. 'I felt disoriented', 'I found it hard to get back to reality').

The researchers additionally indicated that all the modules are to be administered at the end of play sessions as a means for non-intrusive approach. The questionnaire is completed by the players after play sessions which consists of a 5-point unipolar intensity-based answering scales. Authors indicated that the theoretical framework for the questionnaire was performed via partners of the FUGA project while empirical data was gathered via focus group interviews with both hardcore and casual players. It was observed that GEQ has an overall validity based on the semantics of the items. In order to evaluate the set of proposed items, researchers recruited a sizeable sample consisted of 380 participants who played a game of their choice. The survey was performed to test and explore the factor structures of the questionnaires. Researchers indicated that factor

analysis provided structures for the scales that are sensible when compared to theoretical considerations. Additionally, researchers mentioned that the subsequent reliability tests resulted in the construction of reliable subscales with satisfactory to high internal consistencies. Moreover researchers indicated that the GEQ modules have demonstrated discriminant validity since different scales of the survey revealed different response patterns under variations in player type, game content, and setting.

### 3.3.3 Playability Heuristics

As mentioned in the second chapter of this dissertation, Korhonen & Koivisto (2006) have conducted a study to develop a set of heuristics for evaluating player experience of mobile games. They were first to publish a set of heuristics for analyzing player experience on mobile platforms. By reviewing relevant literature that propose heuristics, (Nielsen & Molich, 1990; Nielsen, 1994; Muller et al. 1995; Desurvire et al. 2004; Malone, 1982; Järvinen et al. 2002; Federoff, 2002) researchers have proposed the *Playability Heuristics*. Moreover the heuristic set was consisted of 29 items and 3 modules for evaluating different player experience aspects: *Gameplay*, *Mobility*, and *Game Usability*. Some of the categories and items in the set were originally described in Nokia's Playability Heuristics for Mobile Games (Nokia, 2006). *Gameplay* module incorporates heuristics that cover game mechanics and gameplay issues. The *Game Usability* module covers game controls and the user interface. The module covers aspects similar to usability issues of productivity software but converted to game related problems specifically. Lastly, the *Mobility* module has heuristic items that cover specific issues for mobile games. Additionally, researchers claimed that these modules are common for any mobile game. The *Game Usability* and *Gameplay* modules are indicated as generic modules and be used for evaluating any mobile game regardless of the platform. The researchers attempted to validate the proposed heuristics via conducting expert evaluations using the heuristic set on mobile games. As a result, they indicate that heuristics are useful in identifying playability issues in mobile games as well as helpful to evaluators to focus on important aspects related to the gameplay.

Following their previous work, Korhonen & Koivisto (2007) have conducted another study for defining heuristics for mobile multiplayer games. Because of the modular property of the previously proposed playability heuristic set, they aimed to include multiplayer module for evaluating problems closely related to multiplayer games. By reviewing the literature, the researchers have identified 6 heuristic items and evaluated them via evaluating three multiplayer mobile games. As a result they proposed the *Multiplayer* module with 9 heuristic items.

The playability heuristic were expanded also in 2008 (Korhonen, Saarenpää & Paavilainen, 2008) with the addition of *Context-Aware* module to the playability heuristics set. The researchers have defined this new module for evaluating location based/pervasive mobile games. In order to find pervasive features affecting the player experience, researchers have conducted a user study with six participants. As a result they have defined 4 heuristic items for the module.

Moreover Korhonen (2010) have analyzed differences between player experience evaluations methods by comparing heuristic evaluation to playtest sessions. Utilizing the gameplay and game usability modules of the playability heuristics, he indicated that both methods are complimentary to each other and found several different playability issues. In his dissertation, Korhonen (2016) underline that "*current playability heuristics cover typical problems in the design and they can be applied regardless of the platform or game genre*".

# 4. ANALYZING PLAYABILITY IN MULTI-PLATFORM GAMES: A CASE STUDY OF THE FRUIT NINJA GAME

Technological improvements and novelties have not only provided us with new hardware options including mobile devices, bigger screens and diverse peripherals such as game controllers, but also inevitably led to the adoption of a more user-oriented approach to design more user-friendly platforms. From the first conventional keyboards to the advanced gesture-based interactive technologies, the need for a high-quality user experience has emerged and increased. In the last decade, in order to cope with this transition adopting a user-centered perspective, numerous studies have been conducted making the term user experience (UX) one of the pinnacles of technological evaluation and progression. Video games offer new perspectives for discussions and studies on UX; however, conventional UX and usability methodologies are not sufficient to provide a full understanding of the video game experience, since they mostly focus on productivity in digital platforms rather than entertainment which underpins the gameplay experience. This major difference even leads us to change the terms we use in the context of gaming; replacing ‘usability’ with ‘playability’ and ‘user experience’ (UX) with ‘player experience’ (PX) (Sánchez, 2009; Pagulayan et al. 2003).

PX can be observed in different gaming platforms from consoles to mobile devices and from peripheral-based to gesture-based games requiring novel interaction mechanisms including different screens and peripherals, thus revealing a complex structure at various levels (Pagulayan, 2003). It is critical to understand the complex nature of PX through user research. However, there are only a limited number of studies investigating PX in detail and therefore, this study aims to create a framework for the analysis of this phenomenon based on previous work on usability and UX in the relevant literature.

The available studies on gaming experience only present a set of PX and playability heuristics on a theoretical basis; therefore, there is still a need to conduct empirical research collecting data from various gaming platforms. This study focuses on the analysis of multi-platform player experience based on a PX and playability heuristics framework derived from the literature on usability and UX. The proposed framework is also tested in a multi-platform game setting to seek ways to contribute to the improvement and enrichment of the framework.

### 4.1 Playability and Player Experience

Sánchez et al. (2009a) proposed a new approach to PX with their playability model (PM) revealing the differences of evaluative goals in several specifications. The term playability is used in the analysis of a video game or its design aspects. The researchers identified playability as, “a set of properties that describe the Player Experience using a specific game system whose main objective is to provide enjoyment and entertainment, by being credible and satisfying, when the player plays alone or in company” (Sanchez et al. 2009b).

Several authors (e.g. Sanchez et al. 2009a; Korhonen & Koivisto, 2006; Desurvire et al. 2007; Federoff, 2002; Desurvire et al. 2009) have offered numerous heuristic analysis models to evaluate playability as well as genre-specific models such as heuristics for mobile games (Korhonen & Koivisto, 2006) or advanced table top games (Köffel & Haller, 2008). However, none of these models were validated through tests and/or comparison with other methodologies. We believe that these models need to be validated with user tests in multiple platforms. Usability focuses on utilitarian aspects such as task completion, elimination of errors, external rewards and reducing workload whereas playability is concerned with hedonic aspects such as entertainment, fun to beat obstacles, intrinsic rewards and new things to learn (Lazzaro, 2008). Therefore, conventional usability approaches and heuristics cannot be used alone to describe PX.

The first set of heuristics specific to the structure and model of video games were created by Federoff (2002). To identify playability heuristics, Federoff gathered

traditional usability heuristics from literature review, observed the game development procedures of a company and conducted interviews with the game team of the company. Focusing mostly on game engagement and storyline, Federoff suggested that video games have the following three features; ‘game interface’, ‘game mechanics’ and ‘game playability’.

Another set of heuristics, Heuristics of Playability (PLAY), were proposed by Desurvire (2009) based on the evaluation of usability. The PLAY heuristics were developed with the help of game industry professionals and grouped into three categories: The first category, Gameplay, contained heuristics related to players’ feel of control, challenges, goals, consistencies, balance and the notion of fun; the second was concerned with coolness, entertainment, humor and emotional immersion; and the final category, Usability and Game Mechanics, consisted of the heuristics on documentation, status and score, feedback, terminology, burden on player (as in not putting unnecessary burden on player), screen layout, navigation, error prevention and game story immersion (Desurvire & Wiberg, 2010).

Korhonen and Koivisto (2006) developed another approach to PX, called ‘Playability Heuristics for Mobile Games’, which is closely related to other methodologies regarding playability. The authors proposed a model focusing on mobile technologies and their use in entertainment applications with an emphasis on mobility. They suggested, “The mobile context has some unique characteristics, which require special attention during the evaluation” and discussed these characteristics in detail. Korhonen and Koivisto supported the previous studies in the literature in that traditional usability heuristics lack comprehension and cannot be directly applied.

Even though Korhonen and Koivisto (2006) indicated that usability heuristics cannot be applied to video games, similar to Sánchez et al. (2009) they based their heuristics mostly on user interface. In addition, they pointed out that the former playability approaches were not feasible for mobile platforms and did not offer a novel perspective. Therefore, in their proposed heuristics model, in addition to gameplay and game usability, they included the mobility module. In mobile interaction, users often have environmental

distractions such as lighting, weather conditions or noise. There can also be other people in the vicinity, affecting the gaming experience of users. Therefore, Korhonen and Koivisto identified three main mobility heuristics as ‘the game and play sessions can be started quickly’, ‘the game accommodates with the surroundings’ and ‘interruptions are handled reasonably’ (Korhonen & Koivisto, 2006).

All the above-mentioned methodologies and approaches are viable; however, one study has recently integrated the various definitions and categories of playability heuristics into the framework of ‘playability’. The PM framework developed by Sánchez et al. (2012) provides an easy-to-manage and well standardized set of heuristics to understand PX and its relation with UX. The PM framework is inspired from the three aspects of usability; effectiveness, efficiency and satisfaction as well as their correlation with playability. As a result, seven different heuristics are proposed as follows: effectiveness, learnability, satisfaction, immersion, motivation, emotion and socialization.

## 4.2 Methodology

In this study, the framework of analysis is based on the PM framework of Sánchez et al. (2009, 2012) and the parameters of effectiveness, learnability, satisfaction, immersion, motivation, emotions and socialization. In the tests, we also utilized the mobility heuristics developed by Korhonen and Koivisto (2006).

Since playability analysis is a complex process due to various perspectives, Sánchez et al. suggested “Playability Facets” into consideration, in which they used those facets to categorize different elements of video game architecture. Each facet of playability identify different attributes of playability. The notion of playability facets is to function as a tool to study playability across different video game elements. Along with other methodologies and categorizations, playability facets cover categories of interpersonal and intrapersonal playability.

Mobility heuristics devised by Korhonen and Koivisto (2006), and Interpersonal and Intrapersonal Playability categories proposed by Sánchez et al. are relatively new to the

literature (Figure 4). We acknowledge that they have brought a novel perspective to the playability analysis. In our study, for a better analysis of playability, we introduce a comprehensive approach not only in terms of game features but also concerning the separate evaluation of various complex game elements assessing the parameters in two categories; ‘On-Screen’ and ‘Off-Screen’.

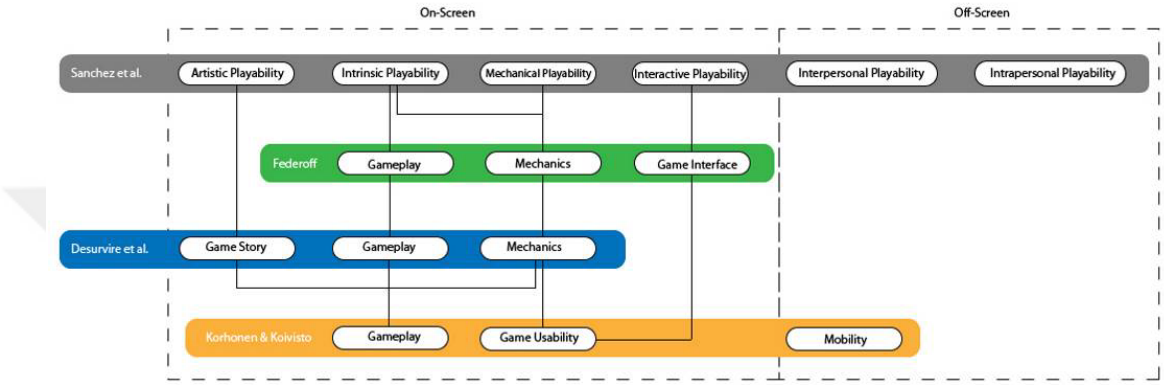



**Figure 2.1** Relations between Video Game Elements for the Analysis of Playability/PX

The notion of “On-Screen” represents the game elements that have been extensively examined in the literature, such as game interface, mechanics, gameplay and storyline. “Off-Screen” represents the aspects that have only been partially explored to date. These are the environmental and social factors affecting the video game experience of individuals.

With the development of cross-platform and multi-platform games, there is a need to consider these social and environmental conditions for a better assessment of PX. The framework we propose allows us to holistically scrutinize all internal and external environmental elements of PX by incorporating the most recent yet incomplete approach by Sánchez et al. and the mobility heuristics developed by Korhonen and Koivisto (2006). In addition, the division between the Off-Screen and On-Screen elements contributes to the comprehensive and complementary structure of this framework. In order to provide an in-depth analysis of multi-platform PX, two different gaming platforms were evaluated; a mobile device (HTC smartphone) and a video game console offering full body gesture-

based interaction (Microsoft Xbox 360). The latter platform was selected for testing, not only because there are limited number of studies regarding the novel interaction possibilities that it provides, but the experience provided might offer physical, mental and social benefits and offer a transition from success oriented play to playing only for fun (Mueller, 2010).

The qualitative differences between these two platforms were examined using the *Fruit Ninja* game which is available for both Xbox 360 and mobile platforms. The F*ruit Ninja* game was originally designed for mobile gaming platforms specifically targeting the mobile phone users. Therefore, the gesture-based version on the Xbox 360 console is only an adapted version of the original game. The game focuses on 'gameplay' rather than other game elements such as narrative. This property has been anticipated to facilitate the comparative data gathering and analysis during the study. Both versions for the platforms had the same game mechanics, in our case using an imaginary sword (finger in mobile version, gestures for the Kinect version) to cut objects. The similarity in gameplay and mechanics enabled us to observe and focus on the experiential differences caused by the gaming platforms.

For the tests, a 50 in. LED TV and Xbox 360 including the Kinect peripheral and a HTC 820 smartphone with a 5.5 in. screen were used. For a more realistic experience, a naturalistic test environment was prepared. Representing one of the major user segments in mobile and console game ecosystem in Turkey, the participants were eight university students from Turkey, dimidiated in terms of gender, with prior experience in using smartphone technologies and ages differing between 19 and 23. Majority of the subjects have played the *Fruit Ninja* game on mobile platform before and only one of them had prior, but limited experience with Kinect peripheral. The test procedure was based on three steps: In the first step, a semi-structured pre-test interview was conducted to collect demographic information about the participants, their gaming background and their experience in relation to using technology and particularly gaming platforms/controllers. The second step was the task observation phase, in which each participant was asked to play the *Fruit Ninja* game on an Android phone and Xbox 360 game console sequentially. We adopted usability testing for understanding and analyzing playability in detail which provided the possibility of witnessing specific quotations from players, which validate

real experiences (Mueller, 2010). With the task flow which reflects a default gaming experience and incorporate essential game attributes so that the players could have realistic experience of the game. Tasks were to enter the game selection screen (task 1), starting the game in classic mode (task 2), going back to the main menu via stopping/pausing the game (task 3), re-entering game selection screen (task 4) and starting and playing the game in arcade mode (task 5). Behavioral data was collected via video recording and mobile eye-tracking equipment. The final step was a post-test interview. This last phase consists of an in-depth interview to understand and analyze player experience in detail. This phase also enabled an attitudinal analysis. The findings of the study were expected to provide an in-depth understanding of PX in two different platforms with a solid framework containing the parameters of usability and UX.

## 4.3 Results and Discussion

### 4.3.1 Effectiveness

The effectiveness of the platforms was evaluated in terms of how easy the game was to play and how much effort it required. All the players stated that the game was easier to control and the tasks were easier to achieve in the mobile platform. Figure 4.1 presents a comparison of the two platforms in terms of the average completion time of tasks for each player (Task 4 was removed from the dataset in the table since it was not included in the Android platform as a new action). The task completion times being less on the mobile device indicates that this platform better facilitated the achievement of specific goals, in our case, the tasks explained in the previous section.

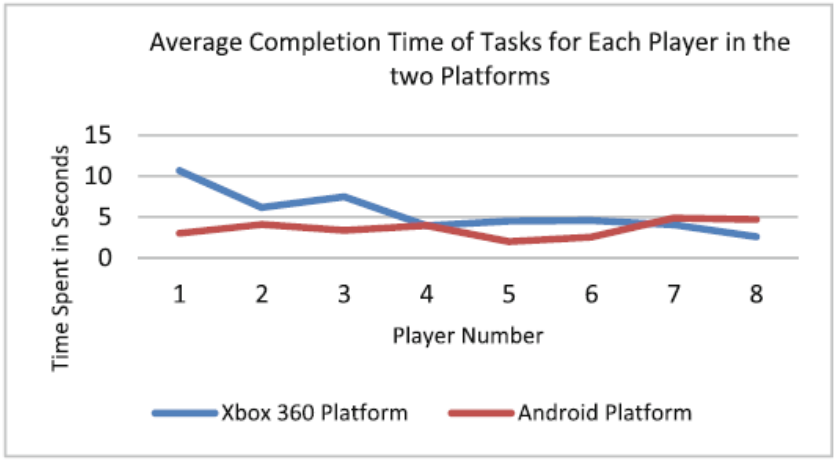



**Figure 4.2** Average Completion Time of Tasks for Each Player in Two Platforms

The third task was about stopping or pausing the game and was not completed by any players in the Xbox 360 platform. This was expected to be found via the trial and error method but none of the actions of the players stopped the game from running. Some players were frustrated and gave responses such as “Do you really believe that we can find a way to exit this game?” (P1), “I think there is no option for it” (P1 & P2), “Would it stop if I turn my back to it?” (P3), and “Stop Kinect stop!” (P4). Participants also gave non-verbal reactions such as turning their back against the screen (P3), waving hands (P3) or even ignoring the task and keeping on playing (P6, P7, P8). Post-tests also showed that the players found the Kinect peripheral much more confusing.

### 4.3.2 Learnability

Learnability refer to the player’s ability to comprehend and master the game system and mechanics. The learnability heuristics are characterized by game knowledge, player skill, difficulty, frustration, speed of learning and discovery. All the participants were able to understand the game mechanics and had no problems or questions about it. The action of slashing fruit came naturally to all players in both platforms (using arms gestures on Xbox and swiping fingers across the screen in the mobile platform). In terms of the player skills, in the Xbox 360 platform, some players experienced peripheral-related problems caused by system feedback. Even when the players were able to react quickly, the system

received the input with a delay, which frustrated the players. However, the players tried to adapt to this defect during play and master the game.

The difficulty aspect of the game can be considered in direct relation to the game completion performance since the *Fruit Ninja* game is mostly based on game mechanics and dynamics. The mobile platform was found to be less difficult to interact with compared to Xbox. The players expressed their frustrations with the difficulty of playing the game as follows: “I needed to cut the fruit with my left hand!” (P1) and “I tried hard, but maybe next time!” (P7). Furthermore, they responded to negative situations by blaming themselves rather than the system although the main reason was the problems with the peripherals.

Regarding the interaction mechanics of the game in the mobile platform, the players mostly acted intuitively and considered it familiar due to prior experience with similar applications. As a result, the participants could quickly relate to the interaction methods. However, in the Xbox 360 platform, the players found the interaction methods to be new, which made space for discovery. During gameplay, some participants tried to improve their skills using different interaction methods such as making a door knocking gesture (P3).

### 4.3.3 Immersion

The immersion heuristics are concerned with conscious awareness, absorption in the game, game realism, control dexterity and socio-cultural proximity with the game. Players acted like ‘ninjas’ mimicking generic martial art movements (such as using the side of their hand to make a cutting gesture to score points). This demonstrated the realism aspect of the system. In other words, the Kinect peripheral provided the players with an opportunity to mimic gestures similar to the observed sociocultural actions. Furthermore, environmental distractions such as other people talking in the room and noise from the street allowed us to observe some players losing their concentration while some others taking no notice of what was going on in their surroundings.

In the study, the players seemed to be under stress when playing the game on Xbox 360. Unlike the mobile platform, many players tended to react anxiously to the loading screens of the Xbox 360 system, biting their nails and/or loosing focus. However, few players responded in the exact opposite way, enjoying the company of the other participants and even dancing during the loading of the game. Therefore, it can be concluded that in the Xbox 360 platform, there is a certain degree of correlation between social capabilities and immersion in the game.

### 4.3.4 Satisfaction

The satisfaction heuristics refer to the gratification or pleasure derived from playing the game. It is characterized by game fun, disappointment with content and attractiveness. During post-tests, the majority of the players stated that they would like to play the game again. Players who preferred Xbox 360 considered this platform to be more fun to interact with due to the novel interaction methods and the bigger screen offered by the system. It was also observed that during gameplay, not only the player but also the other participants had fun since they enjoyed watching or commenting on the active player's effort. This finding may be evaluated in the context of sociability. In terms of aesthetics and content, both platforms offered very similar aspects, rendering it impossible to make a comparison between the two.

### 4.3.5 Motivation

The results regarding motivation give an idea about the game characteristics prompting the player to undertake and continue to perform specific actions until they are completed. The motivation heuristics are characterized by encouragement, curiosity, players' self-improvement and the diversity of game resources. In the current study, six players were eager to play the game again after the tests. The only reason for the loss of players' motivation and encouragement was the errors and unresponsiveness of the system which mostly occurred in the Xbox 360 platform. On the other hand, playing in front of other participants increased the motivation to play the game again.

The players were mostly responsive to the different types of fruit appearing on the screen when playing the game on Xbox 360. The players were curious about how the Kinect peripheral worked, experimenting with several different gestures to achieve the goals and asking other participants about their opinions. However, to play the mobile version of the game, the players only had to swipe their fingers across the screen.

### 4.3.6 Emotion

The emotion heuristics indicate the players' involuntary impulses in response to the stimulus of a video game. The characteristics of emotion are player reaction, game conduct and sensory appeal for game elements. The emotion heuristics are analyzed via observation. The players who played the *Fruit Ninja* game before tended to get more excited when they realized that they could play the game using the Kinect peripheral but after experiencing certain problems with the platform, there was a significant loss of excitement among these players.

In the classic mode of the game, when a big fruit appeared on the screen (as an object to be slashed), the players reacted with enthusiasm, even yelled. One player got so excited that she accidentally pressed the volume button of the mobile phone just after seeing a bigger fruit appearing on the screen. Such actions are examples of impulsive reactions to the given stimulus. In both platforms, the players had fun and control over the game in general and clearly understood the objectives. In the Xbox platform, the participants were more responsive to the actions of the active player and their reactions were also louder when the active player made a mistake.

### 4.3.7 Socialization

The socialization heuristics refer to the set of attributes, elements and resources that promote the social dimensions of the game experience in a group scenario. They have the characteristics of social perception, group awareness, personal implication, sharing of social resources, communication techniques and interaction rules. In this study, the players were in constant social interaction during the tests since they were allowed to watch each

other play and talk during gameplay. This was to observe the social interactions and provide a natural environment to conduct the tests. In addition, we had the chance to explore whether there were any differences between the two platforms in terms of the socialization aspect.

The observations and interviews showed significant differences between the two platforms with Kinect providing a significantly more sociable environment. In the Xbox 360 platform, the players were able to comfortably watch and comment on each other when playing whereas the mobile platform rather isolated the active player. The Xbox 360 platform was observed to encourage group awareness more than the mobile platform. In Xbox 360, at first, the players were competitive but then they quickly shifted to a more cooperative communication resulting in increased personal implication, group awareness and sharing of social resources. Some players shared their ideas about their interaction with Kinect, asking for other participants' help to discover more features of the system. For instance, one player asked the spectators, not the test supervisors, whether he had made a mistake saying, "did I press the wrong button?" (P1). Even though the active and passive players could watch and interact with each other as a group also in the mobile platform, they continued to talk about the Kinect peripheral during the mobile test and how they would try to interact with it next time. In addition, the bigger screen offered by the Xbox 360 platform made it more comfortable for the other players to watch the active player during gameplay.

### 4.3.8 Mobility

In several occasions, in the Xbox 360 platform, the Kinect peripheral needed re-calibration, significantly delaying the start of the game. These calibration problems also caused errors in the system response such as the game starting itself and/or receiving wrong input during gameplay. The length of time the loading screen was displayed was similar in both platforms. Regarding the game accommodation with the surroundings; circumferential elements such as lighting or outside noise affected the game experience in the Xbox 360 platform more than in the mobile platform.

The Kinect peripheral needed to be precisely setup to work flawlessly and even then, caused several problems during the tests since the natural light in the room changed from time to time. Having other players in the vicinity also had an adverse effect on some players, causing them to lose concentration as explained in the immersion heuristics section. However, during the mobile platform tests, the players were also distracted by other players in the vicinity when they heard them talk to each other. This was not particularly a problem for the Xbox 360 platform since all the players watched the game being played and provided feedback to the active player. In terms of interruptions being handled reasonably; the mobile platform was easy to pause and allowed the players to easily go in and out of the game as opposed to the Xbox 360 platform in which none of the players were able to stop or pause the game when asked.

### 4.4 Conclusion

This study analyzed multi-platform PX based on a playability heuristics framework derived from the literature on usability and UX. The proposed framework was tested on the same game in two different platforms to seek ways to contribute to the enrichment and improvement of the existing PX frameworks. Based on the results, it can be concluded that the players were glad that they experienced playing the game on Xbox 360; however, they lost interest in the game due to the playability problems they encountered. At the end of the tests, most players stated that they preferred to play the game on the mobile platform rather than on Xbox 360 with the Kinect peripheral. Despite both platforms offering identical game mechanics, the players had problems reaching certain areas of the screen using the Xbox 360 system due to the bigger screen size. Furthermore, the big screen size negatively affected the players' reaction time to game stimuli. In terms of playability, the mobile platform was more preferable in general but the novelty of the peripheral included in the Xbox 360 platform enhanced the experience. The motivation heuristics were improved since the peripheral encouraged the players to experiment with other possible interaction methods. The results on the immersion heuristics were better for the Kinect peripheral due to its intrinsic quality of rendering the player a direct input and increasing realism going beyond the limitations of the game.

The results on the socialization heuristics indicated that the Xbox 360 platform with a bigger screen size and the Kinect peripheral better enhanced the social experience. However, in terms of effectiveness, the mobile platform players achieved the goals faster and scored better. The Xbox 360 version of the game lacked the necessary tutorials and/or indicators for players regarding some aspects such as pausing the game, thus resulting in frustration. The players stopped interacting with the game from time to time and therefore lost interest in the game. The Kinect peripheral and the bigger screen size in the Xbox platform allowed more social interactions during and even after the tests. In addition, the interaction problems and system errors did not affect the socialization aspect. With its quality of being new in terms of the interaction method used, the Xbox 360 platform caused more excitement at the beginning. The full body gesture-based interaction also enhanced the immersiveness of the game, allowing the players to role-play during the game. However, in terms of mobility, the Kinect peripherals were susceptible to interruptions and the environmental elements were less forgiving. Yet, none of the players were observed to leave the game atmosphere. We consider that games designed specifically for the Kinect peripheral would provide better PX compared to those adapted from the mobile versions.

In the development of the proposed framework, several methodologies were taken into consideration to form a feasible and comprehensive framework to better analyze and understand PX. In this study, we introduced a new framework in which we grouped the game elements into two as Off-Screen and On-Screen elements for a comprehensive and complementary analysis of PX. We found that the mobile heuristics were specifically identified for mobile games and platforms, yet there were limited number of heuristics regarding peripherals such as Kinect and console games that are significantly different in terms of the screen sizes and/or the environment they offer. Future research is needed to identify interchangeable and specific modules of heuristics for specific platforms. There is also the need for comprehensive comparative studies to increase the generalization and validity of our results.

# 5. A MULTI-MODAL APPROACH FOR EVALUATING PLAYER EXPERIENCE ON DIFFERENT GAMING PLATFORMS

Ermi and Mäyrä (2005) defined gaming experience as "an ensemble made up of the player's sensation, thoughts, feelings, actions and meaning-making in gameplay setting" (p. 2). However, the player experience that occurs as a result of all types of interaction with the gaming environment may develop in a positive or negative way depending on the player's perception of entertainment (Kokil & Sanchez, 2015). There are three main factors affecting the quality of player experience; game quality, the quality of the interaction between the player and the game, and the quality of interaction within the context of gaming (e.g., social, temporal, and spatial) (Nacke, Drachen & Göbel, 2010). Research on the analysis and evaluation of all the interactions of the player with the gaming environment during gameplay is crucial (Nordin, Denisova & Cairns, 2014). The understanding of player experience will not only enable to identify the interaction patterns of players in gameplay but will also help various stakeholders from designers to developers in the gaming industry to create games that will meet the needs and expectations of players (Nordin, Denisova & Cairns, 2014).

Today, games are available on various platforms such as desktop or laptop computers, game consoles, mobile smartphones, and tablets. However, in recent years, with the widespread use of smartphones, more players have started to choose mobile games over other platforms (Soomro, Ahmad & Sulaiman, 2013). Since games that were initially developed for desktop computers are now also made available for mobile devices, the gaming platform should be seen as a potential factor that can influence player experience. Despite the necessity to have a multi-modal approach for the determination of players' gaming experience (Poels, Kort & Ijsselsteijn, 2012), there are only few studies in the literature that experimentally investigated player experience on different platforms.

Therefore, the aim of this experimental study was to adopt a multi-modal approach to analyze and evaluate player experience on different platforms, and identify both the factors that potentially have an effect on this experience and the set of methods that are effective on measuring it.

### 5.1 Related Work

In the literature, player experience in games has been investigated by implementing different evaluation methods and employing different evaluation metrics. While some studies have focused on a single game and employed interviews with players to understand their experience, others have utilized multiple games and administered different questionnaires and scales to collect data from players. Poels, Kort and Ijsselsteijn (2012) aimed to understand and categorize gaming experience through semi-structured interviews and individual and group work with 19 players. The authors mainly focused on determining the participants' most notable experience in games, their gameplay motivations and feelings, and overall gaming experience. Lapas and Orehovacki (2015) conducted a survey study to evaluate player experience in interaction with the game environment. Data was collected via an online questionnaire from 158 game players randomly selected from the Massively Online Battle Arena. The authors mainly addressed the player experience elements of immersion, flow, presence, absorption and dissociation, and social play.

In order to investigate player experience in the game environment, some researchers have used recently developed technological devices such as head-up displays or eye trackers. For example, Caroux and Isbister (2016) conducted experimental studies to understand the effects of head-up displays on player experience in two different games; namely, Call of Duty 4: Modern Warfare and Star Craft 2: Wings of Liberty. Data was collected from 15 participants through semi-structured interviews and the eye-movements of players. The authors concluded that not only head-up displays, but also the expertise level of players influenced their gameplay experience. Using an eye-tracking device, Nacke, Stellmach, Sasse and Lindley (2009) evaluated gameplay experience in a gaze-based interactive game environment. To evaluate player experience, the authors collected

data from the self-reports of 30 participants, and their responses to flow and presence questionnaires.

In recent years, questionnaires and scales have been developed to evaluate player experience in games, and researchers have mostly utilized these tools for data collection. Depping et al. (2016) developed a custom game using the Unreal Development Kit to identify the influence of 'disclosing skill assistance' on player experience in games. They used the Player Experience of Needs Satisfaction, Intrinsic Motivation Inventory (IMI), suspense, and attribution scales to collect data from 21 pairs of participants. Employing similar questionnaires and scales, Smeddinck, Mandryk, Birk, Gerling, Barsilowski and Malaka (2016) focused on the effect of game difficulty that can be changed automatically or manually on player experience. In the first part of their study, 40 participants played a game called THYFTHYF that allowed changing the game difficulty manually. After the gameplay session, the participants were asked to complete the Positive and Negative Affect Schedule, Player Experience of Needs Satisfaction, IMI, and Task Load Index. In the second part, the authors repeated the previous study with another casual game, *fl0w*, but this time with different levels of difficulty adjusted automatically.

Gutwin, Rooke, Cockburn, Mandryk and Lafreniere (2016) examined the peak-end effects on player experience in two custom-developed casual games, Match-3 and Whac-A-Germ. Twelve participants were recruited for the study and asked to complete a questionnaire regarding their player experience in games focusing on elements such as perceived performance, game difficulty, perceived accuracy, and player speed. Calvillo-Gamez, Cairns and Cox (2010) tested the Core Elements of the Gaming Questionnaire to understand the gameplay experience of 15 participants, who were asked to play the Tetris game. After the gameplay session, they filled out a questionnaire consisting of several gaming elements, such as enjoyment, frustration, control, ownership, and gameplay.

Korhonen, Montola and Arrasvuori (2009) evaluated the player experience of 13 participants in Grand Theft Auto IV, The Sims 2, and Spore games. The participants were selected from active players and after the gameplay sessions, each participant was interviewed according to the extended version of the Playful Experience framework

including 20 categories, examples of which were captivation, challenge, competition, control, discovery, expression, fantasy, and sensation.

In their first study, Takatalo, Hakkinen, Kaistinen and Nyman (2007) presented two different measurement frameworks to investigate player experience in games. They tested 15 measurement scales, of which eight were extracted for involvement and presence and seven were used to evaluate the flow and quality experience. In a follow-up study, the authors proposed the Presence-Involvement-Flow-Framework (PIFF2) to evaluate the aspects of presence, involvement, and flow in games (Takatalo et al. 2010). PIFF2 was founded on previous research conducted in the field of player experience taking into consideration the basic psychology and the game content. The framework was based upon the use of the Experimental Virtual Environment Experience Questionnaire-Game Pitkä (Takatalo, 2002; Takatalo et al. 2007) and successfully addressed the design problems with the specific sub-components of the game e.g., concerning game mechanics.

Poels, Ijsselsteijn and Kort (2008) introduced the *Game Experience Questionnaire* (GEQ), which offered the most comprehensive framework consisting of self-report measures divided into three different modules; Core, Social Presence, and Post-Game. In this framework, immersion, tension, competence, flow, negative affect, positive affect, and challenge are important factors to offer an understanding and evaluate player experience in games. GEQ has an overall validity based on the semantics of the items; however, there have been only limited number of published studies that adopted a multi-modal approach to evaluate and measure players' gaming experience on different platforms. To the best of our knowledge, in the literature, there is only one study (Kokil and Sánchez, 2015) that examined the effect of different gaming platforms on player experience. In that research, the authors compared PC and touch screen tablet gaming in terms of their impact on player experience. They asked a total of 14 participants to play the *Plants vs. Zombies* game on a PC and tablet in sequence. After each gameplay session, the participants completed the Self-Assessment Manikin, Facets of Playability, PIFF2 and GEQ.

Other studies have only experimented with a single gaming platform to investigate player experience through interviews and GEQ, as well as other related scales fit for the purposes of their research. For example, Al Mahmud, Mubin, Shahid and Martens (2008) focused on the gameplay experience of elderly people and evaluated the design of a tabletop game. They used GEQ as the main questionnaire to collect data from eight volunteering participants. After the completion of the questionnaire, the participants were also interviewed to evaluate their gameplay experience in depth. Oksanen (2013) examined player experience in a collaborative serious game environment. A total of 62 students and 24 teachers were asked to play a multiplayer game, Game Bridge, which allows player collaboration during gameplay. After each gameplay session, the author administered GEQ and the Sociability Scale to understand and evaluate the players' gaming experience.

Urturi, Zapirain and Zorrilla (2015) studied the gameplay experience of elderly people using Microsoft's Kinect sensor. Data was collected from 14 elderly people almost all of whom had previously had very limited gameplay experience. In addition to GEQ, the authors utilized the SUS, which were both filled out by the participants after each gameplay sessions. Similarly, Nacke, Schild and Niesenhaus (2010) analyzed player experience using GEQ in two different games, named Maniac Mansion Deluxe and Zuma games, and collected data from 12 young adults. The authors also administered SUS to understand the correlations between the participants' gameplay experience, reported game quality and usability evaluations.

### 5.2 Methodology

In this experimental study, a mixed-method research design, comprising questionnaires, semi-structured interviews and player observations, was applied to analyze player experience on different gaming platforms. Using a multi-modal approach, player tests were performed both on PCs and mobile devices. GEQ was administered to understand and evaluate player experience on the two different platforms. In addition, in order to better understand player experience, data from GEQ was enriched through semi-structured interviews and player observations. As Nacke, Schild and Niesenhaus (2010)

suggested, GEQ requires the evaluation of the usability dimension of the game and its influence on player experience. Therefore, SUS and the Net Promoter Scale (NPS) were also administered to the participants to obtain data regarding usability. The main objective was to test the appropriateness and applicability of diverse measurement techniques as part of a multi-method approach for gaining an in-depth understanding of player experience.

After gathering the results of player tests, analyses were performed on the data to elucidate the relationships between the variables. The dependent variables were questionnaire, scales, observations and interviews, and the independent variables were screen size, mode of interaction, screen resolution, gaming platforms, and the game. In line with the purposes of the research, the following questions were addressed:

RQ1: How do PC and mobile gaming platforms differ in game experience?

RQ2: What is the relationship between game experience and usability on PC and mobile gaming platforms?

RQ3: How do player observations and interviews contribute to evaluation and understanding of game experience on PC and mobile gaming platforms?

### 5.2.1 Description of the Game

In this study, a casual game, *Plants vs. Zombies*, was used. Initially developed as a PC game for Microsoft Windows operating systems by PopCap Games (http://www.popcap.com) in 2009, the game was later adapted to mobile platforms and made available for iOS operating systems in 2010 and for Android in 2011. The game mechanics, levels, and interface remained the same for all platforms, providing an eligible choice for a comparative analysis of cross platform experience with a focus on the platform rather than the game itself.

The goal of the game is to protect the house from incoming zombies in a garden area with 5 linear lanes. The players are first presented with a variety of plants to choose, which in the game, they place any of these strategically in their garden in any order as they would see fit to defend the area. These plant attack and/or block the zombies automatically when

a zombie enters its lane or square. The game was primarily intended for a single player; yet, there is a multiplayer option in the Xbox console version. However, since the study aimed to compare player experience between PC and mobile platforms, this version was not taken into consideration. Overall, the game has 5 groups of adventure levels, 25 mini games, 20 puzzle mode games, and 11 survival mode levels.

### 5.2.2 Participants

The sample of the study consisted of 20 undergraduate university students (13 male and 7 female). The mean age of the participants was 24.2 (SD = 1.06). All the participants owned a PC and a touch screen smartphone. The participants were selected from casual gamers that had never played the *Plants vs. Zombies* game before. Previous studies on player experience (e.g., Poels, Kort & Ijsselsteijn, 2012) demonstrated that player experience during gameplay may be influenced by several factors such as game genre, player types, player characteristics, and gaming frequencies. Therefore, in order to eliminate the potential effect of such factors on the participants' gameplay experience during the experiment, a purposive sampling method was chosen to select casual gamers. The selection was performed via a demographics survey that also contained questions regarding players' gameplay habits. As a result, only the participants who played games four days or less per week and who had never played *Plants vs. Zombies* were selected**.**

### 5.2.3 Materials

The data on the gameplay experience of the participants was mainly gathered using GEQ, SUS, and NPS. In addition, semi-structured interviews and user observations were also employed. The latter two were utilized to gain a deeper understanding of the overall player experience. GEQ aims to reveal a variety of digital game experiences (Poels, Kort & Ijsselsteijn, 2012; Ijsselsteijn, Poels & Kort, 2013) and is considered to be an effective questionnaire to provide an insight into player experience in the game environment (Nordin, Denisova & Cairns, 2014). GEQ has three different modules covering a range of experiences in a variety of gameplay situations: The core questionnaire, the social presence module and the post-game module. Since the experiment was conducted via a

game that had no networking or multiplayer capabilities and the participants took the test alone without any immediate social influence, the social presence module of GEQ was excluded from the experiment.

The core module of GEQ contains 33 items rated on a five-point scale ranging from 0 (‘not at all’) to 4 (‘extremely’). These items represent 7 categories related to gaming experience; sensory and imaginary immersion, flow, competence, tension/annoyance, challenge, negative affect, and positive affect. The post-game module consists of 17 items to assess player emotions and feelings after a game session. These items represent the four categories of Positive Experience, Negative Experience, Tiredness, and Returning to Reality. In the post-game module, 17 statements were given in order. Similar to the core module, the items in the post-game module of GEQ is rated on a five-point scale ranging from 0 (‘not at all’) to 4 (‘extremely’). Ijsselsteijn, Poels and Kort (2012) pointed out that the post-game module of GEQ was relevant for studies involving naturalistic gaming situations, in which the participant voluntarily starts to play. Since this was not the case in this study, this module was excluded from the analysis.

SUS developed by Brooke (1996) is an instrument to measure the subjective usability of products and services. The scale can be administered quickly and provides parametric scores varying between 0 and 100. The SUS items are rated on a five-point scale from 1 (‘strongly disagree’) to 5 (‘strongly agree’). In accordance with the previous studies that implemented SUS to evaluate gameplay experience (Nacke, Schild & Niesenhaus, 2010; Urturi, Zapirain & Zorrilla, 2015), we slightly altered the scale by replacing the terms ‘system’ with ‘game’, and ‘use’ with ‘play’.

NPS is a metric used to assess the loyalty of customer relationship. It is measured based on responses to a single question proposed by Reichheld (2003) as, “How likely is it that you would recommend our company/product/system/service to a friend or a colleague?” Similar to SUS, the question in NPS was revised as “How likely is it that you would recommend the game to a friend or a colleague?” NPS was added to the test procedures to enrich the findings by exploring the players’ general sense of satisfaction with the game and their tendency to recommend it. The scores range from 0 to 10 and

respondents scoring 0-6 are considered as detractors, 7-8 as passives and 9-10 as promoters (Reichheld & Markey, 2011).

The platforms used for the experiment were an Asus laptop featuring an Intel i3 processor, 4 GB DDR3 RAM, onboard Intel chipset graphics card and 15-inch-wide-screen size and operating on Windows 10 (PC platform) and a Samsung Note 4 smartphone with a 5.7-inch wide screen, Quad 2.7 Ghz processor chipset, 3 GB RAM operating on the Android system version 4.4 (mobile platform). All the gameplay sessions were recorded with a camera. Observation notes were taken by a supervisor during gameplay. The observations and semi structured interviews were voice-recorded for further analysis. Both devices met the required specifications to run the game fluently.

### 5.2.4 Procedure

Experiments were conducted in a game laboratory. The participants were physically isolated from each other during the gameplay sessions to avoid the possibility of any immediate social influence. In order to prevent validity problems related to hardware, all the participants were asked to use the same device provided during the tests. The participants were randomly divided into two groups each consisting of 10 players. One group played the game on the PC platform and the other played it on the mobile platform.

As moderators in the experiment, the researchers introduced themselves, described the *Plants vs. Zombies* game, and explained the main purposes of the experiment before each gameplay session. In the experiments, firstly, each participant was asked to play the game on a given platform (PC or mobile) and took part in an individual gaming session that lasted 15 minutes. During the gameplay sessions, the participants were observed by a supervisor. The observation notes and recordings of the supervisor offered an overall insight into player experience for each participant and a means of identifying the usability problems experienced by the participants.

After the gameplay sessions, the participants were asked to complete GEQ, SUS, and NPS in sequence. Then, a semi-structured interview was conducted to obtain

responses related to the players' experience. Questions were presented regarding the difficulty of the game, whether there was enough time given for play, whether the game was easy to learn, and which specific game features the participants liked or did not like. The interview questions were associated with the statements of GEQ items to understand the responses in depth and to check any possible supporting and consistent relations between the two data collection methods. For a single participant, the completion of the two questionnaires and the interview took approximately 40 minutes in total.

## 5.3 Results

### 5.3.1 Game Experience on the PC and Mobile Platforms

The descriptive measures of central tendency showed that the players tended to be more in favor of PC games compared to mobile games. For instance, those who played the game on the PC platform reported more positive (M = 3.08, SD = 0.81 vs. M = 0.94, SD = 0.26) and less negative affect (M = 0.37, SD = 0.42 vs. M = 2.34, SD = 1.08) than those that played it on the mobile platform. Accordingly, the ratings on negative game experiences such as tension (M = 0.43, SD = 0.49 vs. M = 0.49, SD = 0.47) and negative affect (M = 0.37, SD = 0.42 vs. M = 0.90, SD = 0.92) were higher on the mobile platform compared to the PC platform (Figure 5).

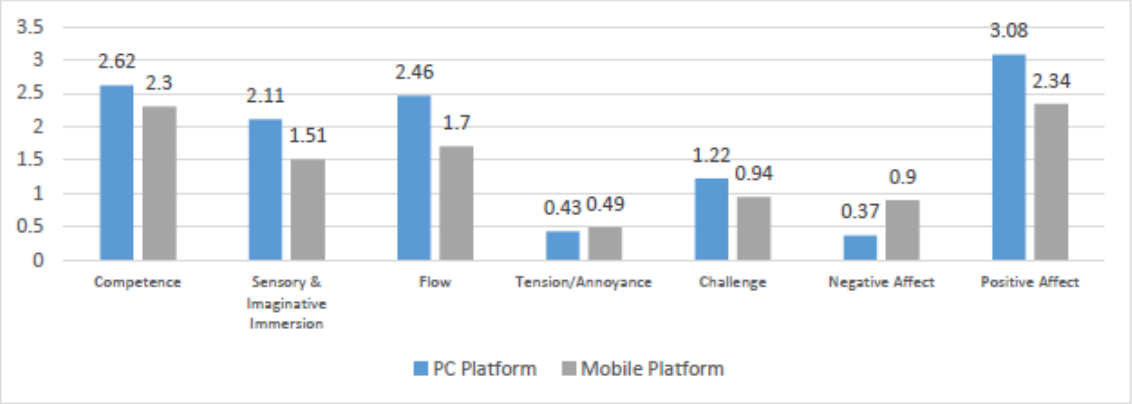



**Figure 5.1** Comparative Game Experience Scores for Each GEQ Item by Game Platform

In order to compare game experiences on the PC and mobile platforms, inter-correlations between the study variables within each platform were examined and compared. Table 5.1 shows the means, standard deviations, and inter-correlations for all game experience items within each game platform; PC and mobile. The intra-platform correlations show that for the PC and mobile platforms, the sensory and imaginative immersion item was positively associated with the flow item ($r = .65$ and $r = .75$, respectively; $p < .05$) and positive affect ($r = .84$ and $r = .92$, respectively; $p < .001$). However, the negative association between the sensory and imaginative immersion item and negative affect was only significant on the mobile platform ($r = -.77$, $p < .001$). On the PC platform, the tension/annoyance item showed a significant association only with

**Table 5.1** Bivariate Correlations between the Game Experience Items for the PC and Mobile Platforms

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Mean | *SD* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Competence | - | .46 | .32 | -.16 | -.56 | -.32 | .52 | .50 | .63 | 2.30 | 0.72 |
| 2. Sensory & Imaginative Immersion | .12 | - | .75* | -.50 | -.28 | -.77** | .92** | .82** | .86** | 1.51 | 0.88 |
| 3. Flow | .07 | .65* | - | -.47 | .01 | -.81** | .77** | .72* | .66* | 1.70 | 0.77 |
| 4. Tension/ Annoyance | .06 | -.52 | -.15 | - | -.20 | .57 | -.55 | -.69* | -.63 | 0.49 | 0.47 |
| 5. Challenge | .10 | .20 | .44 | .21 | - | -.14 | -.15 | -.11 | -.27 | 0.94 | 0.26 |
| 6. Negative Affect | .16 | -.37 | -.23 | .93** | .25 | - | -.90** | -.77** | -.76* | 0.90 | 0.92 |
| 7. Positive Affect | .11 | .84** | .51 | -.66* | .39 | -.56 | - | .79** | .90** | 2.34 | 1.08 |
| 8. SUS | .04 | .89** | .57 | -.58 | .40 | -.43 | .89** | - | .86** | 68.58 | 12.72 |
| 9. NPS | -.06 | .69* | .40 | -.63 | .55 | -.48 | .88** | .91** | - | 5.80 | 2.97 |
| Mean | 2.62 | 2.11 | 2.46 | 0.43 | 1.22 | 0.37 | 3.08 | 79.96 | 6.50 | | |
| *SD* | 0.43 | 0.78 | 1.18 | 0.49 | 0.81 | 0.42 | 0.81 | 11.00 | 3.03 | | |

* $p < .05$, **$p < .001$

negative affect ($r = .93$, $p < .001$) and positive affect ($r = -.66$, $p < .05$). As expected, negative affect was negatively related to positive affect on both platforms although this association was only significant on the mobile platform ($r = -.90$, $p < .001$).

Inter-correlations for the game experience items on the PC and mobile platforms were consistent except for the flow, competence and challenge items. The competence and challenge items on both platforms did show opposite patterns of associations with other study variables, albeit non-significant. For example, the competence item was positively correlated with tension/annoyance, challenge, and negative affect on the PC platform whereas these associations were in a negative direction for the mobile platform. In a similar manner, challenge was positively related to sensory and imaginative immersion and tension/annoyance for the PC players while they emerged as negative for the mobile players. Besides, the flow item was related to negative ($r = -.81$, $p < .001$) and positive affect ($r = .77$, $p < .001$) of the players on the mobile platform while no significant association was observed for the PC players.

On the mobile platform, the players were more positive about coping with the demands of the game, such that they reported less negative affect and tension, whereas on the PC platform, the participants were more intolerant to the challenge of the game because they reported more negative affect and tension. The sense of flow was more informative about and suggestive of affective evaluations of mobile games while flow experience was not related to the affective evaluations of the PC players. Overall, bivariate correlations between the game experience items on the PC and mobile platforms suggested that flow, competence, and challenge were the marker components of game experience, distinguishing the players' game experience on the two platforms.

### 5.3.2 Usability Evaluation of the Game on PC and Mobile Platforms

According to Tullis and Albert (2008), an average SUS of below 60% indicates that the usability of the game is poor and a score of greater than 80% represents good usability. Considering these cut-off values, the players on the PC platform evaluated the game as more usable ($M = 79.9$, $SD = 11.0$) than the players on the mobile platform ($M = 68.6$, SD

= 12.7) (Figure 5.2). Regarding NPS, both the PC and mobile platforms received the same score from the participants. The results showed that on each platform, 3 participants were considered as detractors (scores ranging from 0 to 6), 6 participants were considered as passives (a score of 7 or 8), and 1 participant with a score of 9 was considered as a promoter. Overall, for both platforms, NPS was -20, indicating that the players were not likely to recommend the games to a friend or a colleague.

The SUS and NPS differences between the platforms were examined by analyses of variance (ANOVA). Since we had equal NPS values for both platforms, ANOVA was only performed for SUS. The results showed a statistically significant difference between the SUS values for the two platforms, $F(1, 18) = 4.57$, $p < .05$. The players on the PC platform ($M = 79.95$, $SD = 11.0$) considered the game to be more usable than those that played it on the mobile platform ($M = 68.6$, $SD = 12.7$).

The relationship between SUS, NPS and game experience items showed that on both platforms, the SUS and NPS values were positively associated with positive affect and sensory and imaginative immersion. The relationship between game experience items and SUS and NPS values did differ on the PC and mobile platforms in terms of flow, tension and negative affect items. For instance, on the mobile platform, SUS and NPS were positively related to the flow item and negatively with negative affect. However, these associations were not observed for the PC platform.

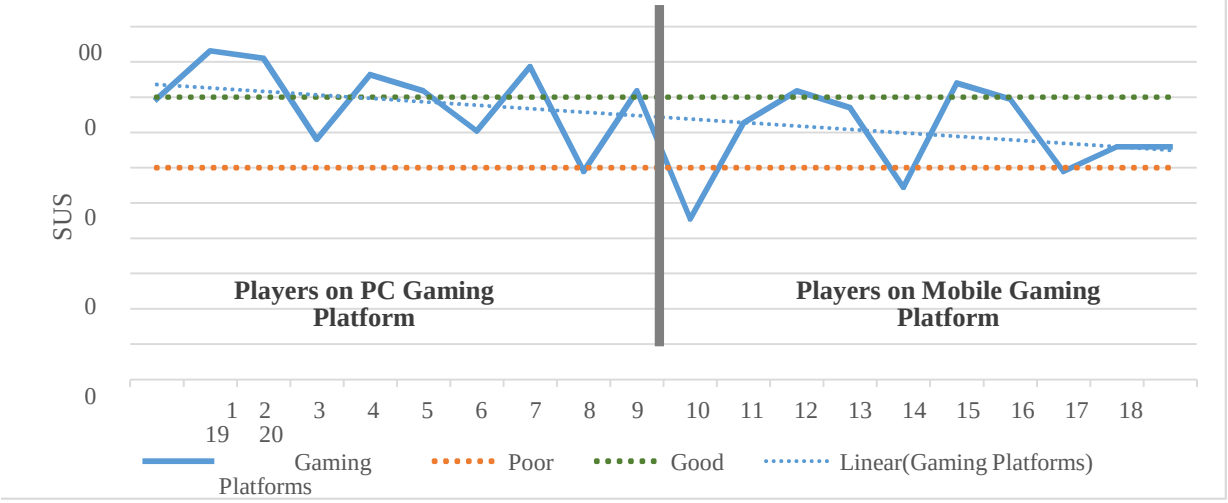



**Figure 5.2** The SUS results for the PC (left) and mobile platforms (right)

In order to identify the contribution of each game experience item to the SUS and NPS evaluations of the players, a series of stepwise regression analyses were conducted. This analysis was chosen because we aimed to explore which game experience item best explained the scores in SUS and NPS. In the analyses, SUS and NPS were treated as dependent variables, and the game experience items (competence, sensory and imaginative immersion, flow, tension/annoyance, challenge, negative affect, and positive affect) were simultaneously entered as independent variables. The stepwise regression analysis was repeated until none of the independent variable contributed to the regression model. According to the results, the game experience items accounted for 79% and 76% of the total variance on SUS and NPS, respectively with positive affect having a notable effect on SUS ($\beta = .89$, $p < .05$) and NPS ($\beta = .88$, $p < .05$) on the PC platform. For the mobile platform, sensory and imaginative immersion ($\beta = .82$, $p < .05$) and positive affect ($\beta = .90$, $p < .05$) significantly contributed to SUS and NPS, accounting for 67% and 80% of the total variance, respectively.

### 5.3.3 Results on Observation and Interviews

Qualitative analysis of the observation notes/recordings and interview answers provided complementary insights for a thorough discussion. Besides, in order to understand the responses in depth and to find out any possible overlapping and consistent relations between the data collected by the GEQ and the interview methods, the results obtained from the interviews were also categorized and grouped by referring to the statements of each game experience items. They showed consistency with the descriptive measures of the central tendency of game experience items in that if a game experience item received a higher score for a platform, the players tended to make more comments associated with the related item. For instance, in terms of the competence item, which include the statements such as “I felt skillful / competent / good at it, etc.”, 18 related keywords such as “*successful*”, “*victorious*” were identified in the interview data from the PC platform players whereas as 16 similar keywords were identified in the answers given by the mobile platform players. Concerning the sensory imaginative and immersion item,

there were 28 relevant keywords for the PC players and 25 keywords for the players on the mobile platform (Figure 5.3).

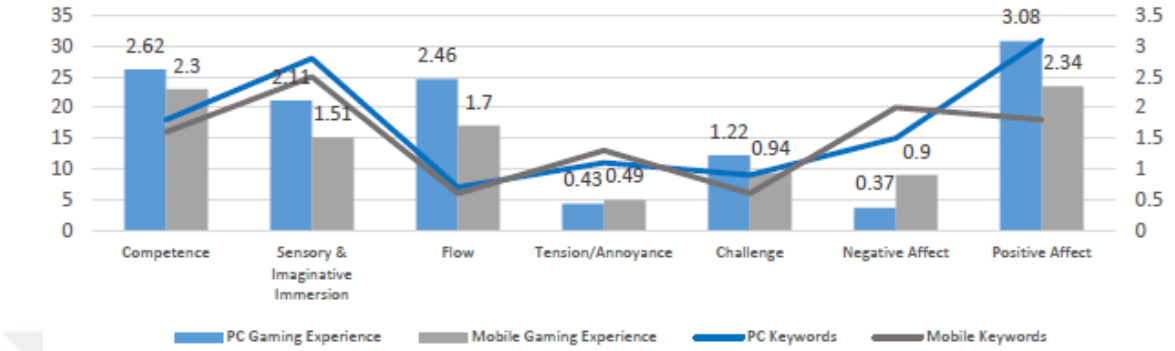



**Figure 5.3** Interview Results and the Mean Score of Game Experience Items

The negative experience items also showed consistency with the players on the PC platform using 11 keywords and those on the mobile platform using 13 keywords for the tension/annoyance item. Similarly, in terms of negative affect, 15 and 20 keywords were identified for the PC and mobile platforms, respectively. The detailed results of observation and interview related to each game experience item are given below.

### 5.3.4 Competence

During the gameplay sessions, the players were observed to have a more positive attitude towards the PC platform and using a mouse rather than interacting with a touch screen. For the PC platform, one player stated, "*I am so used to the mouse peripheral that I can easily have more interactions during the game, more than what they offer*". When asked whether they thought they successfully played the game, 6 and 4 players on the PC and mobile platforms responded positively while only 1 player from each platform felt less successful due to the slow pace of the game. In the interview, when asked if they lost control over the game, 3 mobile platform players stated that they had problems interacting with the device while none of the PC players reported this issue. Four players on the mobile platform considered that the functions of the elements in the game were not sufficiently explained; yet, there was only one person for the PC platform that mentioned this. Besides, 7 players on the PC platform but only one player on the mobile platform

reported that the information cards that explained the function of each plant when the player first earned them at each level were sufficient.

### 5.3.5 Sensory and Imaginative Immersion

The players on the PC platform reported that they were mostly immersed in the game during the gameplay session because of the greater peripheral and sensory stimuli provided by the hardware on this platform. However, the players on the mobile platform were observed to have problems interacting with the mobile device during the gameplay sessions. Most PC players (n = 7) liked the music/sound of the game; yet, the players on the mobile platform not only found the sound element irrelevant for the game but also reported (n = 2) that they would prefer to be able to mute the device/game when playing. However, they did not realize that they could do this through the game options or using the volume buttons. When asked whether it was smooth and easy to play the game on their respective platform, 4 mobile platform players reported that they had problems interacting with the touch screen of the device, of whom 3 also lost control over the game. However, none of the players on the PC platform addressed any problems related to interaction or hardware. The PC players had a more positive attitude toward the game interface with only one player suggesting improvement in the interface while 3 players on the mobile platform made a similar suggestion.

### 5.3.6 Flow

The analysis of the data collected through the researcher's observations and interviews showed that the players on the PC platform did not realize how the time passed during the gameplay session. One player further explained this by stating, "*I would have kept on playing the game because it was getting even more interesting*". The players were also asked to rate the time given for gameplay on a 7 point Likert scale; 1 indicating 'not enough' and 7 'most sufficient'. While the mean score on the PC platform was 2.8, inclined towards 'not enough', the mean score on the mobile platform was 4.1. Within this context, the PC players did not find the gameplay time sufficient and wanted to play keep on playing the game. Concerning the concentration level of the players during the

gameplay, 4 players on the mobile platform reported that they had lost their concentration; yet, only 2 reported this for the PC platform.

### 5.3.7 Tension/Annoyance

The players were asked whether they thought that they had made a mistake during gameplay. As a follow-up question, they were also asked whether they had noticed the availability of a help tool during gameplay. Most players on the PC platform (n = 9) and the mobile platform (n = 8) reported that they were not able to locate and/or get in-game help during gameplay. Concerning the problem of starting the game from the in-game menu, only one player on the PC platform reported that he had experienced problems finding out how to start the game while 6 players on the mobile platform encountered this problem during the gameplay. On the mobile platform, some players asked for the supervisor's help to move further in the test despite no assistance was required for the players on the PC platform.

### 5.3.8 Challenge

One player on the mobile platform stated, "*The game is really slow-paced and easy. Overall, this makes the game really boring*" and another player on the PC platform reported, "*I found the game boring for the entire time I was playing. It was so slow-paced and not challenging; it was a game for people to improve their motor senses*". These responses and observations pointed out the problem concerning the challenge item of the game for both platforms. More than half of the players on both platforms (n = 16) considered the game to be easy to follow. When asked whether they were bored during gameplay, 4 PC players and 8 players on the mobile platform evaluated the general pace of the game as a boring aspect. When asked to rate the challenging aspect of the game on a 7-point Likert scale from 1 'the easiest' to 7 'the hardest', the players scored 1.7 on the PC platform and 2.5 on the mobile platform. This result indicated that the game was considered to be not challenging for both platforms and even perceived to be even easier on the PC platform.

### 5.3.9 Negative Affect

More than half of the players on both platforms (n = 12) evaluated the slow pace of the game as a negative aspect. The players on the PC platform were more eager to have complicated game elements such as a bigger map and more choices of plants (n = 3) while none of the players on the mobile platform commented on this. Similarly, 4 PC players and 1 player on the mobile platform found the plant variety insufficient. Four players on both platforms reported that they would like to have more interaction during the game indicating similar expectations. The players were occasionally observed to be bored due to the slow pace of the game regardless of the platform. One player asked, "*Is this going to continue like this?"* pointing out that the game mechanics were too easy to follow. The players who found the game rather enjoyable tended not to make as many comments during the gameplay.

### 5.3.10 Positive Affect

One player on the mobile platform stated, "*I would play this game on the go; for example, when I am on the bus. The game is easy enough to play on the way*" indicating that the participants' expectations about the game were different for the two platforms. The players were asked about their perceptions of the positive aspects of the game. While 6 PC players found the humor aspect of the game very pleasant, only 2 players on the mobile platform had a similarly positive opinion. Furthermore, in relation to the question whether they would play the game again, 5 players on the mobile platform and 1 PC player had a positive response.

## 5.4 Discussion

It is crucial to acquire an in-depth understanding of player experience in the gaming world, in which both the gaming platforms and game types vary. Though limited in number, there have been attempts to assess player experience on different platforms using different techniques and metrics ranging from heuristics to well-known usability evaluation methods (Aker, Rızvanoğlu, İnal & Yılmaz, 2016). However, none of these

methods provide effective results when applied alone. Poels, Kort and Ijsselsteijn (2012) suggested that gaming experience might be influenced by a number of dynamics such as playing the game alone or with others, player characteristics, game genre, gameplay frequency, gameplay aim, and players' motivation. One of the prominent dynamics that has emerged in recent years is the gaming platform. However, as mentioned in the literature review section, there is a growing need to explore player experience on different gaming platforms by adopting a multi-modal approach. To the best of our knowledge, in the literature, there is only one study (Kokil & Sanchez, 2015) that examined the effect of different gaming platforms on player experience. Therefore, the main objective of this experimental study was to adopt a multi-modal approach to analyze, evaluate, and compare player experience on the PC and mobile gaming platforms.

Concerning how the PC and mobile gaming platforms differ in player experience, inter-correlations between the study variables within each platform were examined and compared with each other. The descriptive measures of central tendency revealed that the participants who played the game on the PC platform reported more positive feelings compared to those on the mobile platform. The mobile platform players also had higher scores in negative items, such as tension/annoyance and negative affect after the gameplay sessions. According to the results, some of the items were associated with each other on both platforms. For example, the sensory and imaginative immersion item was positively associated with the flow and positive affect items on both PC and mobile platforms. This finding is in line with the previous studies reporting that sensory and imaginative immersion was closely related to the item of flow (Oksanen, 2013). This item is, in principal, associated with user engagement and involvement, and involves statements regarding imagination, story, aesthetics, and exploration. The intra-platform correlations also showed that the tension/annoyance item had a significant association only with negative and positive affect. Considering that flow is defined as being dependent on the balance between frustration and boredom, it is reasonable to speculate that as the tension increases, the negative feelings also increase.

The bivariate correlations between the game experience items on each platform showed that flow, competence, and challenge were the marker components of game

experience that differed due to the players' game experience on each platform. Flow is related to a player's ability or competence in relation to playing the game, thus it represents the challenge facing the player (Csikszentmihalyi, 1990). In other words, the flow of the game is determined by the extent of which the player can meet the challenges presented during the gameplay (Csikszentmihalyi, 1996; Kiili & Lainema, 2008). In this context, our results confirm those reported in the literature.

The findings showed that on the PC platform, competence was positively correlated with tension/annoyance, challenge, and negative affect. On the contrary, for the mobile platform, a negatively correlation was observed between competence and these items. Similarly, challenge was positively related to sensory and imaginative immersion and tension/annoyance for the PC players while these correlations were negative for the mobile players. On the PC platform, as the challenge increased with the players' progress through the story and gameplay (increased sensory and imaginative immersion), they started to feel greater tension resulting in negative feelings. On the other hand, challenge seemed to motivate the players on the mobile platform. These players' level of tension decreased as the challenge in the game increased. This finding is consistent with the result that the players on the mobile platform were more positive about coping with the demands of the game and had less negative affect and tension whereas the PC players were not as tolerant to the challenge of the game and consequently, they reported more negative affect and tension. Many researchers have considered flow to be a significant factor affecting player experience (e.g., Inal & Cagiltay, 2007; Admiraal et al. 2011). Similarly, in the present study, the flow item presented as a distinguishing marker specifically for the mobile game experience and it was related to the negative and positive affect of the players on the mobile platform; however, no significant association was observed for the PC players. As the flow increased in the mobile game experience, the positive affect increased accordingly (Figure 5.4).

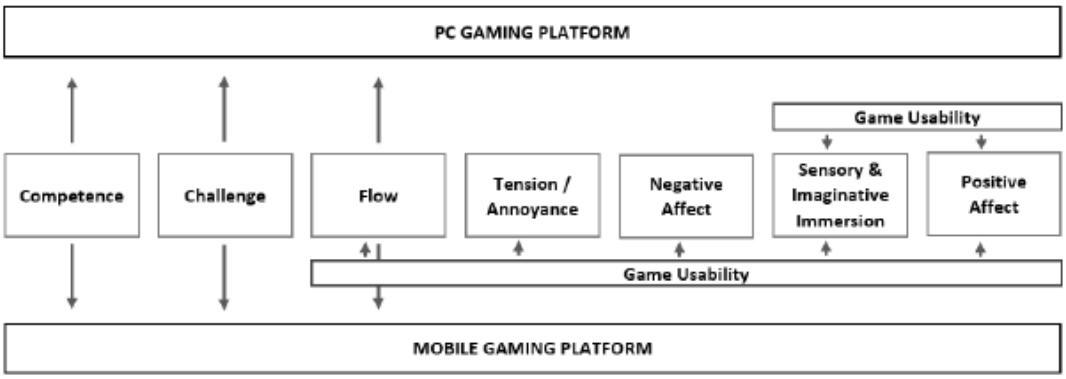



**Figure 5.4** Correlations between Gaming Platforms, Game Experience and Usability Evaluation

Although the players on both platforms were unlikely to recommend the game on both platforms to a friend or a colleague, the players on the PC platform evaluated the game as more usable than those on the mobile platform. Usability score was associated with only one of the marker components of game experience; i.e., flow. Usability scores were also found to be mostly associated with the remaining game experience items that were not considered as marker components. Although usability scores overlapped with the score of only one marker component of game experience, the findings showed that they could still be instrumental in providing complementary findings that contributed to the understanding of player experience.

On both platforms, usability scores were mostly associated with sensory and imaginative immersion and positive affect. Moreover, on the mobile platform, sensory and imaginative immersion and positive affect significantly contributed to usability scores. In this sense, it can be suggested either that when a game was perceived as providing an experience of sensory and imaginative immersion, this affected how the game's usability was rated or that a game with poor usability did not support sensory and imaginative immersion. This finding is partly in agreement with the relevant literature, in which Urturi, Zapirain and Zorrilla (2015) reported no significant correlation between mood rating and usability evaluation; however, Nacke, Schild and Niesenhaus (2010) concluded that the flow and immersion factors were negatively affected by the low

usability evaluation scores of the games. As a contribution, this study showed that the usability scores and game experience items did differ between the PC and mobile platforms in terms of flow, tension, and negative affect.

Usability scores on the mobile platform allowed making better associations between the items compared to the scores on the PC platform. The usability scores were positively related to flow and negatively related to negative affect and tension/annoyance. However, these associations were not observed for the PC platform. It can be suggested that although the mobile platform sometimes caused usability issues due to the limitations of the device and screen, it was still possible to facilitate and maintain flow by presenting challenging tasks and enabling the development of the player's skills. Therefore, it can be concluded that the flow on the mobile platforms increases the players' tolerance for usability issues.

Finally, observation of players during gameplay and in-depth interviews conducted with the players following gameplay sessions proved to be instrumental in offering a detailed insight into the different motivations behind the behaviors and attitudes of the players for both platforms. Moreover, the use of these techniques provided a context to understand the usability issues specific to each problem. The findings on the competence, sensory and imaginative immersion, flow, tension/annoyance, and positive affect items of the game experience were parallel to those derived from the observation notes and the participants' responses during the interviews.

### 5.5 Conclusion

The findings showed that flow, competence, and challenge were the marker components that differed due to the players' game experience on each platform. Although the mobile platform had certain limitations caused by the device and the screen, the increased challenge in the mobile game experience enhanced the feeling of flow and involvement. On the contrary, although the players enjoyed the comfort of playing the game on a big screen that supported detailed graphics and pervasive sound, which contributed to sensory and imaginative immersion, as the challenge increased, the feeling of flow decreased. Usability scores also contributed to this finding. This shows that

providing flow on the mobile platforms allows the players to tolerate usability issues caused by device limitations.

Player observations during gameplay and semi-structured interviews after gameplay sessions demonstrated to be instrumental in offering an in-depth understanding into the different motivations of the players for both PC and mobile platforms. To conclude, it is believed that by covering most aspects of player experience, this study presents a player experience model for the stakeholders of the gaming industry from designers to developers to understand the needs and expectations of players. However, it should be noted that the study was limited in terms of comparing only two gaming platforms with a relatively small sample. Therefore, the presented multi-modal approach should be tested on different gaming platforms and with a larger sample who have different levels of gaming experiences.

# 6. REVISITING HEURISTICS FOR EVALUATING PLAYER EXPERIENCE IN DIFFERENT GAMING PLATFORMS: A MULTI-MODAL APPROACH

Because of their distinctive property of pushing the conventional technology forward, games are considered as one of the most important reasons for novel technologies to be available in our homes and in our everyday lives in the first place (Mayra, 2008). In fact, computer games have become one of the most important and influential sector among the entertainment industry. Human computer interaction studies increasingly focus on computer games and try to provide a guide for the evaluation of games and experiences they offer to players. With the growing influence of computer games in the media and software industry, experts from the field see the importance of analyzing user experience not only for productivity software but also games.

Although games offer a type of user experience, they provide vast and more complex capabilities of interaction with a system. The aim of the software and design considerations as well as complex interaction patterns are fundamentally different than productivity software. Sanchez et al. (2009) stated that usability evaluations are not sufficient and indicated the difference between usability and playability. While user experience research has mostly focused on usability and design topics, there is a lack of understanding concerning how the interactive modalities of games affect user experience (Sutcliffe & Hart, 2017). Better understanding the player experience is necessary not just because it would help identify interaction tendencies of players but also would allow industry designers and developers to produce games that would meet the desires and needs of the players (Nordin, Denisova & Cairns, 2014).

Even though there are many studies conducted for evaluation the user experience (UX) in the literature, studies involving the player experience (PX) are not so many and mostly lack validation. Previously, it was discussed that the use of UX evaluation methods along with PX evaluation methods increases the possibility of analyzing the player experience inclusively (Aker, Inal & Rızvanoğlu, 2017). This multi-modal approach was necessary since the analysis of games with conventional UX methods were mostly insufficient for analyzing the PX.

In today's gaming world, not only PC games rival among each other but consoles and mobile phones are increasingly being opt for gaming. For instance, more players have started choosing mobile platforms over other platforms as a means for gaming (Soomro, Ahmad & Suleiman, 2013). Even though games were formerly developed for PC's at the beginning, they are also made available for other platforms. As is known, these platforms offer different capabilities such as different screen resolution and screen size or possibility to feature novel controls such as tactile feedback. Therefore the gaming platform should be considered to play an important role in terms of player experience. Despite the need of evaluations on player experience, there are only a few studies experimentally inspecting the relations of experiences between gaming platforms (e.g. Zaman et al. 2010; Kokil & Sanchez, 2015; Rafaele et al. 2015; Aker, Inal, Rızvanoğlu, Yılmaz, 2016). These limited number of studies within the literature which investigated player experience on different platforms, often only focus on difference of control mechanisms (e.g. Fritsch et al. 2008; Suhonen & Vaataja 2010; Zaman et al. 2010; Gerling, Klauser & Niesenhaus, 2011).

Previously, the Game Experience Questionnaire (GEQ) (Poels, Kort & Ijsselsteijn, 2012) along with SUS and NPS usability measures, interviews and observations were utilized to analyze experience differences between platforms. This study is a progress that integrates with the previous research (Aker, Inal & Rızvanoğlu 2017). A similar experiment was conducted via employing the Playability Heuristics (Korhonen & Koivisto, 2006 & 2007) to identify and analyze player experience on different platforms. Hence in this study, it was aimed to determine; (1) if heuristic analysis and playability heuristics in specific, is effective for evaluating differences in experience between gaming platforms, (2) whether the usability scales and/or user tests could prove useful for

analyzing platform differences, (3) if the feedbacks from the players could be utilized for analyzing player experience between platforms, (4) if it is possible to provide a holistic point of view by employing a multi-modal approach combining heuristic evaluation with user tests, and lastly (5) which factors and methods are effective for evaluating player experience between gaming platforms.

### 6.1 Related Works

A number of researchers have presented new heuristics for evaluating the gaming experience, and there are several different methodologies utilizing heuristics (e.g. Federoff 2002; Desurvire et al. 2004; Baauw, Bekker & Barendregt, 2005; Röcker & Haar 2006; Korhonen & Koivisto 2006, 2007; Pinelle, et al. 2007; Schaffer, 2007; Pinelle et al. 2008; Jegers 2008; Bernhaupt et al. 2007, 2008). These heuristics are fundamentally based on game design and development areas such as mechanics, interface or gameplay (Federoff, 2002).

Heuristic evaluation is defined as an inspection technique that allows evaluators to examine an interface using statements of usability principles (Nielsen, 1994). This evaluation method is considered to be more effective for evaluating games compared to other methods since this method does not require any task oriented tests and can be employed in a fast and cheap manner (Korhonen, 2010). In this approach, expert-based heuristics evaluation is conducted using simple questions for examining different aspects of the game to find playability problems that may have undesirable effects on the user interaction (Carmody, 2012).

Federoff (2002) have conducted a study on existing heuristics, combining them for evaluating the ‘fun’ of the video games. Expert evaluation was conducted during the study as well as post-game interviews and observations were taken in order to gain additional insights during the tests. After gathering data, Federoff compared the heuristics with Nielsen’s 10 usability heuristics (1994) and proposed a novel set of heuristics for evaluating player experience.

Baauw, Bekker & Barengregt (2005) have conducted another study on utilizing heuristics for evaluating player experience and have developed Structured Expert Evaluation Method. This model was proposed for evaluating children's computer games. Their study mainly aimed for validating the heuristics model. 18 experts were recruited for the tests from the fields of usability and user experience. Four games were evaluated during the tests. Results gathered from expert's notes and heuristics were then compared for analyzing the effectiveness of the set.

Similarly, Sweetser & Wyeth (2005) have proposed a novel set of heuristics, the Game Flow model for evaluating games utilizing the term 'flow' (Csikszentmihalyi, 1990). The proposed set of heuristics were aim to evaluate enjoyment in games. Eight key elements with several heuristic items in each were proposed.

Korhonen & Koivisto (2006) presented the Playability Heuristics. They were first to publish playability heuristics focus on mobile games. They proposed a modular basis for their heuristics, which consisted of game usability, gameplay and mobility. By the means of literature reviews, 29 heuristics were proposed. Some of the categories and heuristics within those categories were developed from Nokia's Playability Heuristics for Mobile Games. Data was collected via expert evaluations with 4 experts, analyzing 5 different mobile games using the heuristics. Subsequently, the set was iteratively improved with the guidance of the experts. Researchers additionally mentioned that the proposed heuristics may be utilized in different platforms and games due to its modular structure. Following their study, Korhonen & Koivisto (2007) published another research in which they included the multiplayer aspects for evaluating mobile games. They presented a multiplayer module for games by employing the heuristics for examining three different games. Nevertheless, Korhonen and Koivisto validated their heuristics with only heuristic evaluations without comparing the results to other methods like playtesting.

Mixed-method evaluation approaches, combining user tests and expert evaluations via heuristics are also common among methods for evaluating player experience and validating the proposed heuristics to some extent. For instance, Desurvire (2004) have proposed the Heuristics of Playability (HEP) and presented 43 heuristic items. These

heuristics were based on literature reviews and evaluated via several experts. The HEP heuristics set contained four categories; gameplay, game story, mechanics and usability. This model was tested on a prototype game. During the study, researchers additionally conducted user-testing methods for comparing the results from both evaluation methods. The user-tests included think-aloud play sessions as well as satisfaction questionnaires. Also the supervisors gathered observation notes. At the end of the study, results from user-tests and heuristic evaluations were compared, indicating that the use of heuristics were efficient for analyzing the player experience.

Gou and Goh (2016) conducted a study where the researchers administered the HEP heuristics (Desurvire, 2004) to evaluate an information literacy game. They have proposed to extend the HEP heuristics framework by including two more categories to the set; characters/graphics and pedagogical effectiveness. In their study, they combined a user-centric approach with the use of heuristics.

Although there is a need for a mixed-method approach for analyzing and determining player experience (Poels, Kort & Ijsselsteijn, 2012), there are only a scarce number of studies within the literature that investigated player experience on different platforms.

Papaloukas et al. (2009) conducted a research with a multi-method approach, combining user tests with expert evaluations. They proposed a modified set of heuristics based on Nielsen's heuristics (1994) for evaluating game usability. Two different games in different platforms were tested (Nintendo Wii & web-based computer game). As for the user tests 30 players were recruited for usability evaluations. Player actions were recorded during the tests and analyzed by three usability experts. Meanwhile, experts played the games for a week and noted possible heuristic items for identifying issues related to the game. Authors resulted their study by indicating the importance of combining these two methods, noting that the use of observations were crucial for enriching the data gathered.

Suhonen & Vaataja (2010) researched the effects of using modular heuristics for health games. During their study, a Nintendo Wii platform game and one mobile game was tested via two experts. Five different heuristic sets were reviewed for using during the tests (Federoff, 2002; Desurvire, 2004; Korhonen & Koivisto, 2006, 2007; Garzotto, 2007) as the authors stated that these heuristic sets were compatible among themselves and complement each other in terms of heuristic items. Later, the researchers stated that Playability Heuristics presented by Korhonen & Koivisto (2006, 2007) was the most eligible for conducting the tests. Also, a similar modular structure was adapted in their study, adding two extra modules for health games; multimodality and persuasiveness.

Similarly, Desurvire & Wiberg (2015) combined different evaluation methods to test the proposed Game Approachability Principles (GAP) heuristic set. They also noted GAP principles held a guiding purpose therefore not directly aimed to evaluate playability. The study took user test results as benchmarks. They utilized usability and heuristics evaluation procedures on different gaming platforms; Xbox 360, PlayStation and Nintendo Wii. Four games from different genres were tested during the study with 32 participants. After the tests, the authors compared and analyzed the results from both approaches. Authors claimed that the heuristics and user-tests supported each other while indicating the best approach to evaluate the general experience of games the use of both methods simultaneously.

### 6.2 Methodology

Despite there is no common agreement on the definition of playability in the literature, researchers have defined the heuristic evaluation method is effective for analyzing the player experience. Reviewing the relevant literature covering heuristics, in terms of development methods, the modular approach by Korhonen and Koivisto is considered to be most valid (Paavilainen, 2010). They proposed a model focused on mobile games while covering gameplay and usability aspects. The proposed heuristics included items derived from previously proposed heuristic sets, generalizable for all platforms as well as a mobile module specifically developed for mobile games. Therefore

to conduct the study, we have employed the Playability Heuristics (Korhonen & Koivisto, 2006).

Following the previously conducted study utilizing the Game Experience Questionnaire (GEQ), this study aims to report a second experiment in which Playability Heuristics are used to evaluate a game to analyze the player experience in different platforms. In this study, a mixed-method approach is employed incorporating questionnaires, semi-structured interviews, observations and playability heuristics to analyze the gaming experience in different platforms. In this multi-modal approach, playability heuristics were administered during the tests. Heuristics were supported with interviews and observations for further information regarding the player experience and comparison between heuristics and playtests. Even though playability heuristics contain the category of 'Game Usability', a modified version of the System Usability Scale (SUS) and Net Promoter Score (NPS) were also administered to inspect possible relations with both our previous study and usability module of the heuristic set. One of the main aims of this study was to examine and test the efficiency of diverse evaluation methods and techniques as a part of the mixed method approach to gain an understanding of player experience on different platforms.

After the tests, we analyzed and compared the results between platforms to observe differences between gaming platforms. The dependent variables were the scores given to heuristics and questionnaires as well as interview & observation feedbacks while independent variables were the gaming platforms in general including properties such as screen size and mode of interaction, and the game. Accordingly, we aimed to inspect if and how the player experience differ between platforms and if it is possible to evaluate these differences via heuristics.

### 6.2.1 The Game

To focus the study on player experience differences between platforms and to be able to evaluate the effectiveness of employing playability heuristics, *Plants vs Zombies* game was chosen as the game to be tested. It was originally developed for PC platform

and for Microsoft Windows operating systems by PopCap Games (http://www.popcap.com/) in 2009. The game was adapted to mobile platforms, iOS operating systems in 2010 and for Android in 2011. The main game elements remained the same between platforms such as mechanics, levels, and interface. This provided a suitable choice for a comparative analysis of player experience between platforms. This essentially allowed a chance to focus on the platform rather than the game itself.

The game has basic interactions and game mechanics, demanding only minimal amount of interaction on both platforms such as clicking or tapping the screen. Players are asked to protect a house from approaching zombies using various plants to stop them. Players are expected to place any of the given plants in their garden in any order as they would see fit to defend the area. When placed, plants react to incoming zombies without any need for further interaction except one plant (sun flower) which is used to provide gaining game currency. Players are expected to click or tap on the sun icons that appear over time to gain currency to plant more. Only the adventure mode of the game was available during the tests since every player that attended the tests had to experience the game from the beginning. However, the game has 5 groups of adventure levels, 25 mini games, 20 puzzle mode games, and 11 survival mode levels, progressively allowing the player to play them.

### 6.2.2 Participants

A total sample of 20 game designers from both game design students and professionals (18 male, 2 female) were recruited for the tests. All subjects had prior gaming experience with both PC and mobile platform. Previous studies on player experience (e.g., Poels, Kort & Ijsselsteijn, 2012) demonstrated that player experience during gameplay may be influenced by several factors such as game genre, player types, player characteristics, and gaming frequencies. Thus, a purposive sampling method was administered for recruiting the players to recruit only the hardcore or mid-core gamers. Also, only the subjects that have never played the game before were chosen to eliminate the factor of familiarity with the game. The selection was performed via a demographics survey that also contained questions regarding players' gaming habits. As a result, only

the participants who played games four days or more per week and who had never played *Plants vs. Zombies* were selected. Korhonen & Koivisto (2006) stated that the evaluators should have at least some amount of game design expertise. As a result, only the subjects with a game design and/or user experience knowledge were chosen to explore expert review method and the use of heuristics for evaluating player experience.

### 6.2.3 Material

The data for the experiment was gathered mainly from the playability heuristics, SUS and NPS surveys. Moreover, interviews and user observations were employed to gather further and deeper understanding of the overall player experience.

Playability heuristics are proposed to overcome some missing elements in previously proposed heuristics such as dealing with issues related to mobile platform and overlapping definitions which made them ambiguous according as the authors state in their study (Korhonen & Koivisto, 2006). In their research, they proposed heuristics to evaluate mobile games via three modules; “Gameplay”, “Game Usability” and “Mobility” (2006). These modules were discussed and compared to other heuristics during recent years, providing a more refined and extensive version. In this study, we used this recent version of Playability Heuristics from (Korhonen, 2016). This heuristic set consisted 5 main modules: “Context-aware”, “Multiplayer”, “Gameplay”, “Game Usability” and “Mobility” and 47 heuristics. Since our aim is to inspect and analyze differences of player experience between PC and mobile games, we utilized only the latter three modules. Accordingly, Gameplay module consists of 14 items and is claimed to be valid across all platforms, while the Game Usability module has 12 items and was proposed to cover the issues regarding game controls and interface as well as including common usability issues. Lastly the mobility module consists of 7 items. It contains heuristics that are specific for mobile games. During the tests, the playability heuristic items were rated on a five-point Likert scale to make the evaluation process feasible for the evaluators. On the scale, answers varied from 1 (strongly agree) to 5 (strongly disagree). During the tests, players were asked to comment freely if they wanted to make comments to potentially gather

insights about he answers for the heuristic set. Additionally, post-test interviews were made after the play sessions aimed to analyze the answers given to the heuristic set.

As mentioned before, SUS and NPS was administered after the heuristic evaluations during the tests. The System Usability Scale (SUS) provides a quick tool for measuring the usability. It is a survey to measure the subjective usability of products and services, developed by Brooke (1996). It consists of a 10 item questionnaire with five response options for respondents; from ‘strongly agree’ to ‘strongly disagree’. The scale provides parametric scores varying between 0 and 100. Similar to previous researches (Nacke, Schild & Niesenhaus, 2010; Urturi, Zapirain & Zorrilla, 2015), to employ the survey for evaluating games, a slightly modified version of the test was administered by replacing the terms ‘system’ with ‘game’, and ‘use’ with ‘play’.

Fundamentally Net Promoter Score (NPS) is a tool for measuring customer experience. It is used essentially to assess the loyalty of the customer. This metric was included to the procedure to strengthen the analysis of the player experience for assessing players’ general sense of satisfaction and tendency to recommend it. It consists of a single question using a 0-10 scale; “How likely is it that you would recommend our company/product/system/service to a friend or a colleague?” as proposed by Reichheld (2003). Like the modified version of the SUS, the question was revised as “How likely is it that you would recommend the game to a friend or a colleague?” Respondents scoring 0-6 are grouped as detractors, 7-8 as passives and 9-10 as promoters (Reichheld & Markey, 2011).

For the PC platform, an *Asus* laptop featuring an Intel i3 processor, 4 GB DDR3 RAM, onboard Intel chipset graphics card and 15-inch-wide-screen size and operating on Windows 10 was used. For the mobile platform, a Samsung Note 5 mobile phone featuring an Octa-core (4x2.1 GHz Cortex-A57 & 4x1.5 GHz Cortex-A53) processor, 4 GB RAM and a 5.7 screen and operating on Android version 7.0 was used. Both of the hardware specifications of the platforms were sufficient for running the game fluently. All the participants were asked to use the same device provided during the tests to prevent validity problems that may relate to hardware capabilities.

### 6.2.4 Procedure

Experiments were conducted in a focus group laboratory with audio and video recording capability. The participants were randomly divided into two groups each consisting of 10 players. They were admitted to the test separately and one by one to avoid any bias and social influence. One group played the game on the PC platform and the other played it on the mobile platform.

One researcher moderated the test sessions. The researcher introduced himself, described the *Plants vs. Zombies* game, and explained the main purposes of the experiment before each gameplay session. The moderator have attended the tests directly and took observation notes as well as player comments. At the beginning of the tests, each participant was asked to use the given platform (PC or mobile) and took part in an individual gaming session that lasted as much as the player wanted. Observation notes of the supervisor and the recordings provided an overall insight into player experience for each participant and a means of identifying the usability problems experienced by the participants. The subjects were asked to fill the playability heuristics set during the tests whether during the gameplay sessions or after finishing playing the game, allowing an appropriate condition for the subjects to not feel any time pressure while conducting an expert evaluation.

After completing the heuristic set, the subjects were asked to fill the SUS and NPS questionnaires followed by an interview respectively. The interview questions were associated with the statements of Playability Heuristics to understand the responses in depth and to check any possible supporting and consistent relations between the two data collection methods. All the tests, including the interviews took two hours on average per player.

## 6.3 Results

Before conducting analysis of the results, it was expected to receive scores which would indicate platform differences via game usability module of the heuristic set as well

as the SUS scores. Since Korhonen & Koivisto (2006) implied that the Gameplay and Game Usability modules were applicable to all the games and Game Usability module was more focused on the usability issues of the games, we have expected to receive different scores between platforms mainly from the Game Usability module. Nevertheless, the results were analyzed and compared in order to point out any possible implication to assess the difference between platforms in terms of playability.

After conducting the tests, we first analyzed which heuristics were pointing out a difference between platforms. In order to inspect differences between answers for the heuristics set, an independent-samples t-test was conducted to compare heuristic scores between PC and mobile platforms. The total mean of the heuristic scores did not indicate a significant difference between mobile (M= 3.81, SD=0.47) and PC (M= 4.05, SD=0.40) platforms, yet indicated that PC players perceived less playability problems compared to mobile players. Moreover, we compared the two heuristic module means between platforms. According to this, our results indicated there was not a significant difference in the scores for Game Usability module for mobile (M=3.96, SD=0.14) and PC (M= 4.18, SD=0.40) platforms, indicating less usability problems for PC platform. Similarly, the scores for Gameplay module did not indicate a significant difference between mobile (M=3.67, SD=0.60) and PC (M=3.94, SD=053) platforms but again indicated that the players favored PC platform over mobile in terms of gameplay.

Although the heuristic set means did not indicate any difference in terms of significance, when inspected on the scale of items, only the Game Usability heuristic, *GU6* (“Navigation is consistent, logical, and minimalist”*)* have shown a statistically significant difference between mobile (M=3.1, S=0.87) and PC(M=4.4, S=0.84) platforms t(18)=-3.38, P=0,003. For analysis, descriptive statistics (Table 6 & 6.1) were derived from the test scores.

**Table 6.1** Descriptive Statistics for Playability Heuristics Game Usability & Gameplay Modules

| ***Playability Heuristics Items*** | **Mobile** | | **PC** | | **Overall** | |
|---|---|---|---|---|---|---|
| | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| **GU1a** *"Audio-visual representation supports the game"* | 4 | .94 | 4.7 | .48 | 4.35 | .81 |
| **GU1b** *"A view to the game-world supports smooth interaction and the camera behaves correctly"* | 4.4 | .69 | 4.5 | .52 | 4.45 | .60 |
| **GU2** *"Screen layout is efficient and visually pleasing"* | 3.7 | 1.05 | 4.2 | .63 | 3.95 | .88 |
| **GU3** *"Device UI and game UI are used for their own purposes"* | 4.2 | .91 | 4.6 | .51 | 4.4 | .75 |
| **GU4** *"Indicators are visible"* | 3.8 | 1.03 | 3.5 | .97 | 3.65 | .98 |
| **GU5** *"The player understands the terminology"* | 3.9 | 1.10 | 4.4 | .51 | 4.15 | .8 |
| **GU6** *"Navigation is consistent, logical, and minimalist"* | 3.1 | .87 | 4.4 | .84 | 3.75 | 1.06 |
| **GU7** *"Game controllers are consistent and follow standard conventions"* | 4.6 | .51 | 4.5 | .70 | 4.55 | .60 |
| **GU8** *"Game controls are convenient and flexible"* | 4.1 | .87 | 4.2 | 1.22 | 4.15 | 1.03 |
| **GU9** *"The game gives feedback on the player's actions"* | 4.3 | .67 | 4.3 | .82 | 4.3 | .73 |
| **GU10** *"The player cannot make irreversible errors"* | 3.4 | 1.57 | 3.4 | 1.34 | 3.4 | 1.42 |
| **GU11** *"The player does not have to memorize things unnecessarily"* | 4.5 | .97 | 4 | .94 | 4.25 | .96 |
| **GU12** *"The game contains help"* | 3.5 | .97 | 3.7 | .94 | 3.6 | .94 |
| **GP1** *"The game provides clear goals or supports player-created goals"* | 4.2 | 1.03 | 4.6 | .69 | 4.4 | .88 |
| **GP2** *"The player sees the progress in the game and can compare the results"* | 3.7 | .94 | 3.9 | .99 | 3.8 | .95 |
| **GP3** *"The players are rewarded and the rewards are meaningful"* | 3.6 | 1.17 | 3.7 | 1.15 | 3.65 | 1.13 |
| **GP4** *"Player is in control"* | 4.6 | .51 | 4.4 | .69 | 4.5 | .60 |
| **GP5** *"Challenge, strategy, and pace are in balance"* | 3.2 | 1.54 | 4 | 1.15 | 3.6 | 1.39 |
| **GP6** *"The first-time experience is encouraging"* | 4.1 | .99 | 4.4 | .69 | 4.25 | .85 |
| **GP7** *"The game story, if any, supports the gameplay and is meaningful"* | 3.5 | .97 | 3.9 | .73 | 3.7 | .86 |
| **GP8** *"There are no repetitive or boring tasks"* | 3.1 | 1.28 | 2.6 | .84 | 2.85 | 1.08 |
| **GP9** *"The players can express themselves"* | 2.5 | 1.50 | 3.3 | 1.25 | 2.9 | 1.41 |
| **GP10** *"The game supports different playing"* | 3.4 | 1.50 | 3.7 | 1.49 | 3.55 | 1.46 |
| **GP11** *"The game does not stagnate"* | 3.4 | .96 | 3.6 | .96 | 3.5 | .94 |
| **GP12** *"The game is consistent"* | 3.8 | 1.03 | 4.5 | .52 | 4.15 | .87 |
| **GP13** *"The game uses orthogonal unit differentiation"* | 4.1 | .87 | 4.5 | .52 | 4.30 | .73 |
| **GP14** *"The player does not lose any hard-won possessions"* | 4.2 | .91 | 4 | 1.41 | 4.10 | 1.16 |

**Table 6.2** Descriptive Statistics for Playability Heuristics Mobility Module

| ***Mobility Heuristics Module Items (MO)*** | ***M*** | ***SD*** |
|---|---|---|
| **MO1** *"The play sessions can be started quickly"* | 3,80 | 1,22 |
| **MO2** *"The game accommodates the surroundings"* | 4,70 | ,48 |
| **MO3** *"Interruptions are handled reasonably"* | 4,00 | ,47 |
| **MO4** *"The graphical design is accommodated to current brightness (Supplements GU1a)"* | 3,90 | 1,19 |
| **MO5** *"The player should be aware of some device features while playing (Supplements GU3 and GU4)"* | 3,00 | 1,24 |
| **MO6** *"Mobile devices have their own conventions for input (Supplements GU7)"* | 3,90 | ,73 |
| **MO7** *"The tutorial should respond to immediate demand (Supplements GU12)"* | 3,70 | 1,41 |

## 6.4 Results of Observations & Interviews

Karat (1994) indicates that the number of identified usability issues is one of the subjects that define the effectiveness of an evaluation method when comparing diverse evaluation methods. Hence, most of the studies that utilize heuristic evaluation along with playtests have conducted analysis based on the number of playability problems discovered whether by the heuristics or the playtests (e.g. Desurvire et al. 2004; Korhonen, 2006, 2010). Nevertheless, Desurvire et al. (2004) noted that playtests have additionally provided results indicating specific problems that are not mentioned in the heuristics and are crucial for evaluating games. Moreover, Korhonen (2016) indicates; "*One characteristic of the game evaluations is to think about the origin of the playability problem.*" Therefore, it is important to identify the origins of the problems that we received from the interviews and observations as well as identifying specific issues that could not directly associated with the heuristic set.

A similar approach from Hara & Ovasaka (2014) was followed for grouping the interview results and observations to further identify problems indicated during the tests. Although their study involves identification of problems from game reviews from online sources via utilizing keywords, a similar approach through gathered interview and observation results facilitated the analysis process. Similar to this, Soomro et al. (2012)

have conducted a preliminary research on playability heuristics for mobile games by developing problem categories to group the problems identified from the interviews. Instead of utilizing keywords from the feedbacks, an in-depth analysis on the causes of the problems was conducted for this study. This grouping of problems is also aimed to facilitate identification of heuristics which would be useful for evaluating player experience between platforms during the study.

To further evaluate the results, qualitative analysis of the interviews and observation notes provided supportive results for a thorough discussion. Moreover, it was aimed to identify possible overlapping responses and relations between heuristics in which the participants may have indicated consistencies pointing towards the platform differences. For instance it was identified that some heuristics could not be affiliated directly to the platform capabilities while some can directly be related to game elements. By this grouping of problems, it was possible to put forward a general classification of differences in experience between platforms.

During our study, each heuristic is found to be rooted in a specific problem. Because of these diversity of specific problems related to both platforms and the game tested, we defined different sets of problem groupings based on the feedbacks and observations. Interview statements of the players and observations were identified jointly for each heuristic item (Appendix-A) for the betterment of qualitative analysis.

### 6.4.1 Grouping of Problems

As Federoff (2002) mentioned, the proposed heuristics are generally and fundamentally based on game design and development areas such as mechanics, interface or gameplay and may not be directly related with the causes of playability problems. Furthermore, current heuristic evaluations still rely heavily on similar elements to define the heuristic categories. Although this might be useful for dissecting the playability problems, it was clear that a novel approach was necessary for analyzing experience specifically in different platforms. Moreover, playability heuristics were not directly compatible with platform related problems that are indicated during the tests. Therefore,

descriptions of problem categories were mapped with the causes of the indicated problems and interrelated with heuristic items. As whole, we identified 3 different problem groups regarding the dependencies of the heuristics based on results from the feedbacks in our analysis: *Design-Dependent Problems*, *Game-Specific Problems* & *Device-Dependent Problems*. These categories were defined to provide insights about the issues that participants have indicated about the game and possible solutions to the problems encountered in the context of specific problem areas.

### 6.4.2 Device-Dependent Problems

The device-dependent problems consist of responses to heuristic items and feedbacks which indicated issues related to the gaming device. Ergonomic issues and/or other hardware related problems were included in this set of category (see Appendix-A for the details of the results). The results received from the players during interviews and observations indicated that, the problem category consists of issues related to device capabilities such as the size of the screen, interaction mode, peripherals or quality of the speakers. These problems are indicated to refer hardware differences and/or proposed to be in direct relation with device characteristics.

### 6.4.3 Design-Dependent Problems

The design-dependent problems include feedbacks and heuristic items depending to game mechanics and interactions such as conflicting / overlapping interaction mechanics between the game and the device operating system. These issues were defined in order to provide insights for the game designers and developers (See Appendix-A for the results received from the players during interviews and observations). The design-dependent problems are examined to be related with the game elements and include issues such as interface problems or gameplay issues. This category fundamentally address platform-specific problems that can be resolved by better development/design of the game and game elements specifically crafted for the platform.

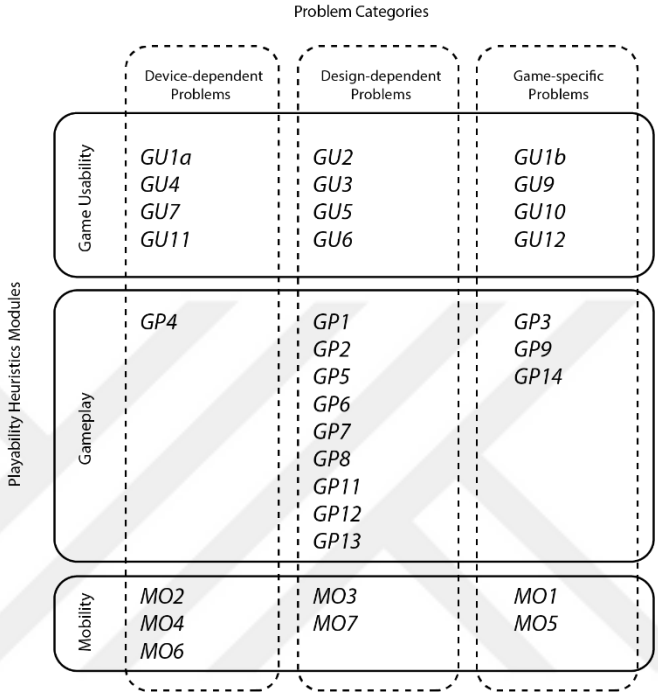
### 6.4.4 Game-Specific Problems

It has been indicated that different genres have different problems (Pinelle, 2008). The game-specific problems include heuristic items and feedbacks received from participants related directly to the particular game genre and specific game elements. The items which were included in this category are found to be not applicable for our research since they did not provide evidence regarding the differences between platforms. (See Appendix-A for the results received from the players during interviews and observations). Game-specific problems are examined to be related with issues related with the game genre in particular. Heuristic items and feedbacks related to this category are indicated as either non-existent in the game or not useful for observing platform differences.

## 6.5 Discussion

In order to set the analysis out in full, each result of the heuristic item and their problem groupings are evaluated in detail in the following section. The groupings of the heuristics and related observations allowed for further analysis of each heuristic item from the perspective of the cause of the playability problem. Based on this approach, heuristic items are discussed in accordance to their grouping category (Table 6.2). It should be noted that the mobility heuristic module did not include heuristics that are derived from the notion of differences between platforms. Hence the heuristic items do not directly address differences between platforms in terms of experience but underline some of the issues that players have encountered on mobile platform. All the mobile players scored the module items above the average in general which indicated that they were content with the game within the range of given heuristics. Moreover the mobility heuristic items are not applicable in general to indicate any platform difference between platforms. Nevertheless, associations with other heuristic modules were inspected and relations were addressed during the analysis.

**Table 6.3** Problem Categories Associated With Playability Heuristics




### 6.5.1 Analysis of Device-Dependent Problems

As mentioned above, these problems and associated heuristics described device related issues such as ergonomics or hardware capabilities. 8 playability heuristic items (*GU1a*, *GU4*, *GU7*, *GU11*, *GP4*, *MO2*, *MO4* and *MO6*) are found to be in association with this problem category.

The results from the interviews and observations indicated that mobile players have mostly commented negatively on the sound effects of the game. All PC players (n=10) gave positive comments on audio-visual representations including keywords like "*I liked this, good, well-matched*", while in mobile platform several players mentioned keywords such as "*bad, didn't like, annoying*" (n=3) to indicate that they did not like the music and the sound effects. Three players (PM4, PM5 & PM8) mentioned that although they liked the graphics in general, they were annoyed by the sound effects of the game during play. Additionally, two players on the mobile platform have accidentally blocked the speaker

of the phone with their hands because of the holding position, resulting in a severely muffled sound during play. On the contrary, none of the PC players (n=10) indicated any complaints regarding sounds or graphics of the game and were observed to be content with the audio-visual aspects of the game. These problems are identified to be related with the playability heuristic item *GU1a* ("Audio-visual representation supports the game"). Results pointed out that the audio and the graphics could be considered as separate heuristics (for example under a sub-section). The complaints were related mostly to the limited speaker capabilities of the mobile platform and the placement of the speakers on the device. Therefore it may be considered as a hardware issue in general.

As Wood et al. (2005) suggested, direct interaction on the display requires less cognitive, spatial or attentional demand. Moreover, computers may cause unnecessary physical and cognitive loads depending on the hardware or software (Laux, 2001). Because of these suggestions, mobile players are inspected to perceive indicators easily compared to PC players. In the literature, it is suggested that small screen sizes can actually be more effective since it reduces the visual load (Nattkemper & Prinz, 1990). The presentation of few menu items at a time could to be helpful in general, because it is widely known that the legibility and also the readability is hampered by increased density of text on the screen (e.g. Norman, 1991; Ziefle et al. 2005). Thus, Ziefle (2010) said; "*From this it can be deduced that visually demanding displays negatively influence information access*". According to the test results and feedbacks, it is observed that players' perception for indicators can potentially be positively affected by the screen size of the mobile platform. Therefore the heuristic item *GU4* ("Indicators are visible") is considered to be relevant for the problem category. In terms of memorization of game related elements and the cognitive load, depending on the feedbacks and interviews, the heuristic item *GU11* is found to be in close relationship with the heuristic item *GU4* and the ease of use of the mobile platform. Detenber & Reeves (1996) suggested that there is a close correlation between bigger screen size and memorization. Even so, the results indicate that mobile players gave higher scores to the heuristic item *GU11* ("The player does not have to memorize things unnecessarily"). Observation notes indicated that two PC players (PP6, PP3) mentioned negative comments on the subject and gave lower scores to the related heuristic accordingly. Both players indicated that they were not able to see

the info about the given plants whenever they wanted to and got frustrated about using them. Mobile players experienced the same game design in this manner but they all pointed out that the game was giving them the chance to learn by trial and error. In general, it was examined that players' perceptions and cognitive demands of the hardware varies between platforms, favoring mobile platform.

Although the game offers the same in-game mechanisms, the fundamental differences between controls were observed to be not addressed via heuristics. Essentially, mobile platform offered the capability to manipulate the object on the screen directly by either touching or dragging during the game while PC platform only allowed the players to interact with a mouse via clicking to the objects. Thus, the differences in controls were considered to hold crucial potential for identifying platform differences yet we could not observe any differences on interviews and/or observations. In terms of heuristics, a difference between average scores for the heuristic items *GU7* ("Game controllers are consistent and follow standard conventions") & *GU8* ("Game controls are convenient and flexible") was expected since the platforms offer fundamentally different interaction methods and peripherals as one is limited to touch screen and one to a mouse. For the heuristic item *GU7*, the average scores for mobile players were higher when compared to PC players by only a narrow margin. As for the latter heuristic item, *GU8,* PC players gave slightly higher scores on average. These somewhat similar scores indicate that all the players from both platforms found the controls acceptable, consistent and convenient in general, even though these heuristics items imply the notion of controls capable of being flexible for different preferences of players. The game did not provide such an option for re-mapping any of the controls. It was also noticed that the wordings of these heuristics did not relate to the fundamental control differences between platforms in terms of usability. For instance a mobile player (PM9) mentioned that he would like to have drag & drop type of interaction yet the game already provided an additional drag and drop mechanism for the mobile platform. Even so, the same mobile player gave a high score to both of the heuristics. Additionally, the heuristic item *MO6* ("Mobile devices have their own conventions for input") was described as in supplementary relationship with the heuristic item *GU7*. This association is validated through observations. The players which gave higher scores indicated that the conventions for the mobile platform were obvious

and natural. While the players gave higher scores for the item on average (M=3.90, SD=0.73), two players (PM1 & PM10) gave low scores, indicating that they were feeling indecisive. It was observed that these two players had difficulty understanding the heuristic item *MO6* and hesitated to answer either positively or negatively. Similarly, mobile players indicated that they felt more in control during the game when answering the heuristic item *GP4* ("Player is in control"). This was observed to be in parallel with the touch screen interaction requiring less training and hand-eye coordination, and being easier to use effectively because of the reduced cognitive load (Thomas & Milan, 1987). Although players on mobile platform had several problems interacting with the game such as tapping on the wrong plants, they indicated that they felt in control throughout the game. This contrast between errors and players feeling of direct control on mobile platform may become more significant compared to the context of usability and cause a higher feeling of control hence more error tolerance.

In terms of mobile device capabilities for adapting to environmental factors such as loud noises coming from outside or low brightness conditions, players gave almost the highest score regarding the heuristic item *MO2* ("The game accommodates the surroundings") pointing out that they were comfortable adjusting some of the game features while playing the game without being affected from the surroundings. Analysis of the interviews and observations indicated that players were able to adjust the volume easily during the play via in-game options menu (n=7) or by using the devices general volume button (n=3). In relation to that, interviews indicated that the players understood the heuristic item as referring to device capabilities rather than the game itself. Similar to these findings, analysis of observations and interviews indicated that *MO4* ("The graphical design is accommodated to current brightness") is strongly related with the heuristic item *MO2* in terms of given context. Korhonen (2016) also indicated that this heuristic item is given as a supplement to the heuristic *GU1a,* yet when analyzed, the results suggested that the players did not relate the item directly with *GU1a* but with the device capabilities.

### 6.5.2 Analysis of Design-Dependent Problems

The design-dependent problems are problems that relate to issues regarding interface of gameplay which essentially address platform-specific problems which can be resolved by developmental precautions and/or modifications. 15 playability heuristic items (*GU2, GU3, GU5, GU6, GP1, GP2, GP5, GP6, GP7, GP8, GP11, GP12* and *GP13*) are found to be in association with this problem category. The association with most of the gameplay module heuristics was anticipated since these heuristic items also relate mostly to game design & development issues. Different that the other problem categories, suggestions were given during the analysis of the problems in this category to indicate probable solutions.

Results from observations and interviews indicated that none of the players from both platforms mentioned negative comments related to the aesthetic properties of the layout. Nevertheless, the scores for the heuristic item *GU2* ("Screen layout is efficient and visually pleasing") indicated that the overall screen layout was found to be more pleasing by PC players. During the interviews, all PC players (n=10) used positive keywords such as "*easy, appropriate, clear*" regarding the layout, yet only 2 of them mentioned a negative aspect related to the screen resolution of the game being not high-definition. Mobile players have mentioned negative keywords such as "*didn't like, tiring*" (n=3). Mobile players also have encountered accidents related with the interaction of the device (n=5). These accidents are observed to be directly related to the operating system of the mobile device, specifically the *Android* operating systems' menu button being accidentally pressed during play. Furthermore, in terms of layout of the in-game elements, there is a significant difference between two versions of the game. On mobile platform, the progress bar of the game is presented on top of the screen while and on PC, it is placed at the bottom of the screen. One player on PC platform (PP8) indicated that he was fully aware of the progress bar while one mobile player (PM5) highlighted a problem of him not being able to identify on which level he was playing at. Lauer (1979) said, "*When everything is emphasized, nothing is emphasized.*" Too many focal points are likely to confuse players and may diffuse their interest. It was identified that on the mobile platform, additional interactions focus on top of the screen area, such as game-related

active objects dropping from the top of the screen. The menu button is also located on top of the screen. On the contrary, the progress bar on the PC platform is placed at the bottom without any extra icons or objects nearby, isolating and therefore enhancing the visibility of the progress bar during the game. Because of these feedbacks and observations, we identified this problem in association with *GP2* ("The player sees the progress in the game and can compare results"). The results of the observations suggest that progress indicators should be designed according to the platform. Indicators should also provide sufficient and convenient information of the overall progress of the game as well as short term progress.

Similar with the heuristic *GU2*, the observations and interviews highlighted that fundamentally the difference in average scores of *GU3* ("Device UI and game UI are used for their own purposes"), are related to the interaction with the operating system. Both platforms players start the game in the default full screen mode. Accordingly, the heuristic item *GU3* suggests the game should be presented in full screen mode to better immerse the player. However, the accidental press of the *Android* menu button drops the player out of the game, potentially breaking the feeling of immersion. Because of its context, this heuristic item is also considered as one of the crucial heuristic items to indicate a difference between platforms. In contrast, although most of the mobile players have accidentally tapped the *Android* menu button during play, the scores given to the *MO3* ("Interruptions are handled reasonably") heuristic item was high. Players were observed to accidentally press the device menu button. Afterwards, they get back to the game from where they were interrupted immediately by pressing the same button again. They mentioned that they were able to return to the game in a fast manner.

The results indicated that mobile players have more difficulty on receiving information in terms of terminology including explanations of plants and the game story. Although several PC players gave negative comments (n=3) regarding the terminology by using keywords such as "*didn't understand, not enough, complex*", mostly mobile players have encountered problems regarding the functions of the plants during our observations. Since the problem is mainly encountered because of the terminological issues, this problem is examined to be related to the heuristic item *GU5* ("The player understands the

terminology"). It was also observed that mobile players had the tendency to skip important information screens (such as pop-up boxes giving extra information about specific functions of the game) during the tutorials in a very fast manner. As a result many mobile players had problems with understanding the exact functions of the plants (n=7). Additionally, three mobile players also tried to find a way to re-check the plant information without success (PM1, PM6, and PM7). Exceptionally one mobile player indicated that even after reading the explanations of the plants, he did not receive sufficient information (PM5). Moreover none of the PC players have observed to be skipping the information screens providing plant information during the play. Furthermore, both player groups' made similar mistakes during the game by misinterpreting a function of a plant. Nevertheless mobile players complained mostly on how the game was not sufficiently explaining the functions of the given plants during the game and related this with the consistency of the game to evaluate the relation between plant information and functions of the plants. Players from both platforms understood that the terminological problem was also related to the heuristic item *GP12* ("The game is consistent"). Lastly, results from the feedbacks indicated that the players on mobile platform had difficulty in understanding specific functions and strategic implications of some of the plants. This specific problem of not utilizing plants in a strategic fashion was identified to be in relation with *GP13* ("The game uses orthogonal unit differentiation"). To give an example, PM7 experienced similar problems utilizing specific plants during the game and was not able to employ a plants specific capabilities during the game. The player tried using a specific plant several times during the play without success. It should be noted that the function of this plant was explained in the information screens, pointing out its use. Still, PM7 has skipped this information screen during the play. Developing the game primarily based on the targeted platform capabilities and player tendencies towards the platform would increase positive feedback from the players. Since the mobile players had the tendency to skip text-based information screens during play, a mobile specific version of the information should be implemented, specifically developed for the mobile version of the game such as animated explanations and/or small tutorial videos.

The limited screen space of the mobile platform is a problematic issue for providing optimized information access and navigation (Zhao et al., 2001). The results from the

observations and the interviews indicated the notion of the problematic navigation on the mobile platform. PC players mostly commented positively on the navigation of the game, using keywords such as; "*clear, easy, plain*" (n=8) while mobile players mentioned mostly negative comments using keywords like "*bad, couldn't find, problematic*" (n=6). Because of the context of the feedbacks, the heuristic item *GU6* ("Navigation is consistent, logical and minimalist") is associated with the problem. As explained on the previous section of the study, this heuristic item indicated a significant difference between platforms according to the average scores given to the heuristic. Ziefle (2010) suggested that: "*Field independence, the ability to separate an item from the context of which it is a part, exerts a strong influence upon navigation ability*". Primary observations indicated that mobile players had severe stagnation in the main menu of the game while trying to find the function button to start playing. Players were observed to be unable to perceive the buttons separate from the context from the game background in the main menu. For both versions of the game, the main menu (or the start menu) screen included a tombstone graphic with a button embedded on it which is mapped to the function of starting the game. Additionally, on the PC version of the game the start button on the main menu screen was highlighted when the mouse pointer crossed over it. Even though PC players did not have any issues finding the buttons, mobile players had a hard time finding the embedded button to start the game. During the tests, mobile players indicated that they couldn't immediately understand how they would begin playing the game since they were not able to directly locate the start button. Two mobile platform players were not able to find how to start the game asked for help from the supervisor. The results suggest that on mobile, navigation paths and design elements related to navigation of the game should be designed specifically for mobile platforms. The general progression throughout the game also should be designed in consideration of mobile platform, providing as clear, functional and basic information as possible.

It was identified that the onboarding process works rather better than on the mobile platform. We observed that the tutorials at the beginning of the game provided sufficient information related to the mechanics yet did not provide information of long-term goals of the game. The results regarding the heuristic item *GP1* ("The game provides clear goals or supports player-created goals") indicated that players on PC platform understood the

necessary goals and ways to achieve them more clearly. It was also observed that the familiarity with the *tower defense* game genre may had an impact regarding the answers. While most of the players were familiar with the game genre, two mobile players (PM6, PM9) indicated that they have never played the game genre on a mobile phone before. Moreover, PC players felt more incentive while playing the game at the beginning. This may indicate that the heuristic item *GP6* ("The first-time experience is encouraging") is also related to the onboarding process which is mentioned for the heuristic *GP1*. It was noted that first-time experience for mobile players was problematic because of the start menu problems. Accordingly, a mobile-first approach can be suggested for a more beneficial evaluation of the item. These findings suggest that the game developers and designers should consider designing the explanations of the goals and the onboarding processes specific to mobile platforms to provide better experiences. A step-by-step development approach for the mobile platform would potentially more effective and provide a better onboarding experience for mobile players.

Another crucial factor affecting the player experience is definitely the challenge and the pacing of the game. Most of the players from both platforms complained about the pacing of the game and its lack of challenge, yet mobile players especially commented on this during observations. While all mobile platform players gave negative comments using keywords such as; "*dull, too easy, boring*" (n=10), several PC players liked the challenge and the pace of the game by giving comments using keywords like; "*tireless, stress-free*" (n=3). Therefore we can indicate that the PC version of the game is perceived to be less cumbersome compared to the mobile platform. These feedbacks are directly related with the heuristic item *GP5* ("Challenge, strategy and pace are in balance") because of its context. In terms of the pacing of the game, one player playing the game on PC platform had a negative comment about being bored during play (PP6), while several mobile players mentioned this problem by using keywords such as "boring, slow" (n=3). Additionally, mobile players mentioned that they were familiar to the tower defense game genre for mobile platform, but the game lacked some of their expectations. These feedbacks are found to be in relation to the heuristic item *GP8* ("There are no repetitive or boring tasks"). Results received regarding the heuristic item suggest that the tasks during the game were perceived to be more boring and/or repetitive on the PC platform.

Controversially, feedbacks regarding the heuristic item *GP11* (“The game does not stagnate”) indicate that PC players felt less stagnation and more progress during the play sessions. Moreover, it was observed that most of the players from both platforms associated this heuristic item to the pacing of the game. As indicated in the *GP5* heuristic item analysis, all of the mobile players complained about the slow pace of the game during play sessions. It is known that frequent achievements could provide a better experience in general on mobile platform yet the game was ported from PC to mobile in the first place. The game genre is considered as suitable for the mobile platform by mobile players yet it lacked some of the important game elements such as sufficient variations of game-related tasks. It was understood that the developers may have to design/port mobile versions of the games with renewed challenge and pacing mechanisms. The results suggested that if the game was designed with the notion of mobile platform and specifically offered more variations of tasks during the game, players would enjoy the game more.

Regarding the story of the game, PC players indicated that the story of the game felt fun and enjoyable while some mobile players mentioned that they did not understand the story of the game and not found the ambient crucial for the concept of the game (n=2). The onboarding process differences between platforms is once more observed regarding this feedback. Moreover, because of its context, the heuristic item *GP7* (“The game story, if any, supports the gameplay and is meaningful”) is found to be in direct association with the feedbacks. According to the scores received from the heuristic item, mobile players had difficulty understanding the overall story of the game and purpose of game characters in terms of context. The item is also observed to be associated with *GU5*, suggesting that mobile players have more difficulty on receiving information in terms of terminology including explanations of plants and the game story. The developers and designers of the game might enhance the overall experience by integrating a story specifically crafted for the platform.

Regardless of the fact that there is no immediate and/or on-demand tutorial function in the game, players gave relatively higher scores to the *MO7* (“The tutorial should respond to immediate demand”) heuristic item. Players mentioned they were content regarding the heuristic item except one player (PM4). When analyzed in-depth,

observations suggested that most of the players affiliated the heuristic with the tutorial levels instead of any immediate help function. The item is referred to *GU12* heuristic item in the proposed playability heuristics (Korhonen, 2016) yet for instance PM6 gave the lowest score to *MO7* while giving the highest score to *GU12* heuristic item. These results underline that the proposed links between these heuristics might not be there when applied in practice since the game fundamentally does not offer on demand help or tutorial. Players would benefit from having immediate tutorial when they needed assistance related to the functions of the plants.

### 6.5.3 Analysis of Game-Specific Problems

The game-specific problems are identified as the problems that relate to issues which are particular to the game genre and/or specific game elements that do not allow evaluating experience differences between platforms. 8 playability heuristic items (*GU1b, GU9, GU10, GU12, GP3, GP9, GP14, MO1,* and *MO5*) are found to be in association with this category.

As mentioned above, it should be noted that the game does not directly provide a help function neither for PC nor mobile platform. However the game provides well designed tutorial levels to the players. Additionally nearly all of the players mentioned that they did not need any help function during the game, sometimes referring to lack of challenge of the game. The observations and interviews indicated that the players were relating the heuristic item *GU12* ("The game contains help") to the lack of help functionality during the game yet the heuristic item also explained to include tutorials of games. Because of these reasons, it is possible to suggest that an extra heuristic item following this heuristic should be implemented to the set (such as *GU12b*), separating the role of tutorials and the concept of learnability from the heuristic item. Nevertheless, because of the current state of the heuristic item, it was identified as not useful for evaluating platform differences since some games fundamentally do not necessitate a help function at all.

In terms of *GU1b* (“A view to the game-world supports smooth interaction and the camera behaves correctly”), the game did not have any dynamic camera interaction mechanism. Therefore the heuristic is identified as to be suggested to a specific type of a game genre such as third-person shooters. Similarly, the heuristic items *GU9* (“The game gives feedback on the player’s actions”) and *GU10* (“The player cannot make irreversible errors”) were found to be inefficient for evaluating differences between platforms since these two heuristics are about internal dynamics of the game which are considered to be out of the context for analyzing differences between platforms.

In terms of rewards given during play, the reward mechanics in the game was developed in the same manner for both of the platforms. The game basically offers new plants after the player completes each level in both versions. Other than the progression related rewards, there was no score indicator. Because of these reasons, no difference between platforms in the context of this heuristic was expected. The reward mechanics were related to the heuristic item *GP3* (“The players are rewarded and the rewards are meaningful”) in which no notable difference was observed regarding the heuristic item scores. In relation with the reward mechanics, the heuristic item *GP14* (“The player does not lose any hard-won possessions”) was identified as not applicable for analyzing differences between platforms. This was also previously expected before the tests since the game for both platforms have the same mechanics and dynamics and estimated to fulfill this heuristic sufficiently.

In terms of the heuristic item *GP9* (“The players can express themselves” & *GP10* (“The game supports different playing styles”), the game does not allow players to customize their avatars or personal preferences. Although, PC players mentioned the personalization of different strategies during the play sessions might count as expressions. During observations, several players said that modification of the game world could be considered relatable with this heuristic (n=3 for PC & n=2 for mobile), yet none of the players were conclusive about this comment. Despite these suggestions, observations indicated that this heuristic item does not apply for neither of the versions of the game and is not applicable to evaluate platform differences in terms of player experience. Similarly, *GP10* (“The game supports different playing styles”) did not receive any feedback

regarding the difference between platforms since the game did not offer any capability to choose a role and/or style of play in the world.

The opening speed of the game was somewhat similar between platforms. The results of the scores for the heuristic item *MO1* ("The play sessions can be started quickly") indicated that mobile players were content with the opening speed of the game sessions. One player (PM9) was observed to tap on the screen while the game was first loading on the device. When asked, the player mentioned that it was not annoying and he tapped only because he was curious to see what would happen. It was observed that most of the mobile players were referring this heuristic to the relation between device capabilities and game requirements, such as more demanding games require more time to start. Thus it could be suggested that the item would be useful for comparing games in mobile platforms rather than platforms.

Lastly, the heuristic item *MO5* ("The player should be aware of some device features while playing") was found to be controversial among several players. Observations indicated that players did not relate this item with neither of those heuristic items and considered it to be not related with the game at all. A number of mobile players (n=3) gave fairly low scores to this item, suggesting that the item was not viable. This item was also proposed as supplementary for the items *GU3* and *GU4* (Korhonen, 2016). When compared, several players who gave low scores to this item did not give scores to the supplementing *GU3* and *GU4* items in the same fashion.

### 6.5.4 Usability Evaluation of the Game

Regarding the NPS scores, PC players scored higher (10) when compared to mobile players (-50). The results showed that, 2 players were considered as detractors (scores ranging from 0 to 6) while 5 players were considered as passives (a score of 7 or 8), and 3 players were considered as promoters (scores higher than 8). As for the mobile platform, 5 players were considered as detractors. Meanwhile the remaining 5 players were considered as passives. According to the NPS scores, no promoters are identified in the mobile player group. These results indicate a gap between platforms in terms of player

tendencies of recommending the game. Mobile players are not likely to recommend the game to their friends or colleagues.

Tullis and Albert (2008), suggested that an average SUS of below 60% indicate poor usability of the game and a score of greater than 80% represents good usability. According to this, players on PC platform evaluated the game as more usable (M=83.5, S=1.63) when compared to mobile players (M=76.2, SD=7.47) (Table 6.3).

**Table 6.4** Descriptive Statistics of SUS Scores

| ***SUS Questions*** | **Mobile** | | **PC** | | **Overall** | |
|---|---|---|---|---|---|---|
| | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** |
| **Q1** *"I think that I would like to play this game frequently"* | 2.3 | 1.05 | 3.5 | 1.17 | 2.9 | 1.25 |
| **Q2** *"I found the game unnecessarily complex"* | 1.5 | .70 | 1.2 | .42 | 1.35 | .58 |
| **Q3** *"I thought the game was easy to play"* | 4.6 | .51 | 4.8 | .42 | 4.7 | .47 |
| **Q4** *"I think that I would need the support of a technical person to be able to play this game"* | 1.1 | .31 | 1 | 0 | 1.05 | .22 |
| **Q5** *"I found that the various functions in this game were well integrated"* | 3.7 | .67 | 3.8 | .63 | 3.75 | .63 |
| **Q6** *"I thought that there was too much inconsistency in this game"* | 1.5 | .70 | 1.5 | .52 | 1.5 | .60 |
| **Q7** *"I would imagine that most people would learn to play this game very quickly"* | 4.7 | .48 | 4.6 | .51 | 4.65 | .48 |
| **Q8** *"I found the game very cumbersome to use"* | 3.3 | 1.15 | 3.1 | .87 | 3.2 | 1.0 |
| **Q9** *"I felt very confident playing the game"* | 4.3 | .67 | 4.7 | .48 | 4.5 | .60 |
| **Q10** *"I needed to learn a lot of things before I could get going with this game"* | 1.6 | .96 | 1.2 | .42 | 1.4 | .75 |

The SUS and NPS differences between platforms were inspected by independent samples t-tests. After conducting an independent samples t-test for SUS scores between platforms, SUS tests indicated a significant difference between mobile and PC platforms; t(18)=-2.52, p=0.21. Furthermore NPS scores indicated that there was not a significant difference between PC (M=7.1, SD=2.6) and mobile platform (M=5.8, SD=2.2) conditions; t(18)=-1.18, p=.252.

These results suggest that NPS survey helped identify player tendency to recommend the game in general meanwhile SUS tests indicate platform differences in terms of usability. When analyzed in detail, three SUS questions (1st, 9th & 10th) were found to be in relation with the observations. It should also be noted that these SUS item scores were derived from average scores from the players and not normalized. This proved to be beneficial for comparing the results directly with the context of the items. For the betterment of the qualitative analysis, these SUS questions were analyzed in detail (Table 6.3).

***SUS1: "I think that I would like to play this game frequently"***

PC players gave higher scores to this item (M=3.5, SD=1.17) when compared to mobile players (M=2.3, SD=1.05). PC players mentioned that they were not bothered by the pace of the game and not felt bored as much as mobile players throughout the game. As mentioned in the heuristic *GP5,* one of the PC players (PP10) indicated that they felt "*stress-free*". Additionally, when observed via heuristic results, PC platform can be seen as more comfortable rather than mobile. Additionally, Net Promoter Scores (NPS) were parallel with this results. NPS score for PC players was 10 while mobile players scored -50.

***SUS9: "I felt very confident playing the game"***

PC players gave higher scores regarding this SUS item (M=4.7, SD=0.48) when compared to mobile players (M=4.3, SD=0.67). When the inter-correlations between SUS items and heuristics were examined, mobile players have shown significant association between the SUS item with *GU11* ($r = .762$, $p < .05$) while PC platform players have shown significant association between the SUS item with *GU7* ($r = .813$, $p < .001$), *GU8* ($r = .674$, $p < .05$), *GU12* ($r = .752$, $p < .05$), *GP3* ($r = .813$, $p < .001$), *GP13* ($r = .655$, $p < .05$). These results suggested that PC players associated several game aspects related to the SUS item while mobile players mainly focused on only one heuristic. Apart from that, PC players gave more positive comments related to the SUS item when compared to mobile players.

***SUS10: "I needed to learn a lot of things before I could get going with this game"***

Mobile players gave higher scores to this SUS item (M=1.6, SD=0.96) compared to PC players (M=1.2, SD=0.42). When the feedbacks from the players and observations were analyzed, PC players indicated more tendency for understanding the game as can be seen on the heuristic item. Analysis suggests that several scopes of the given heuristics such as terminology, onboarding, ease of use and game story are affecting the SUS item.

## 6.6 Conclusion

It is crucial to understand the importance of player experience in the gaming world. Fundamentally, platform variations may be crucial in terms of experience since different platforms offer different experiences. There are only a number of studies examining player experience using various methodologies such as utilization of heuristics, employing metrics and/or surveys. But none of those methods has proven to be effective and sufficiently holistic when applied merely by themselves and do not cover all the aspects regarding player experience. Because of these reasons, we have conducted a multi-modal study utilizing heuristics and playtests for evaluating player experience in different platforms. This multi-modality in our approach proved to be effective in terms of getting detailed insights and analysis as well as further improving the evaluation. Additionally, interviews and player observations after gameplay sessions were contributive in offering an in-depth comprehension of different motivations of the players for both PC and mobile platforms.

Korhonen & Koivisto (2006) indicated that playability heuristics would be effective for evaluating the player experience. Moreover, they suggested that the modular structure of the heuristic set can involve most of the features of games such as Game Usability module covering game controls and interfaces as well as containing common usability aspects that help players. The results from our study indicated that the heuristic set was ineffective for analyzing different experiences between platforms. Detailed analysis of the feedbacks suggested that the heuristic set needed more improvements.

We found that wordings of several heuristic items were not easily comprehensible by players and resulted in misunderstandings. Most importantly, we identified that players do not relate heuristic items directly with their given module, such as several gameplay module items were associated with game usability issues and vice-versa. Due to feedbacks and observations we received during the tests, we were able to identify three distinctive problem groups from the feedbacks of the players and were able to associate them with the heuristics. These problem categories could potentially be useful for analyzing playtests in combination with heuristics from a structured perspective when evaluating platform differences. This grouping of problems is also proved to be useful for identification of heuristics which would be useful for evaluating player experience between platforms during the study. We identified that the heuristic items which are included in the *design-dependent problems* category could be analyzed by focusing on development stages of games specifically with the platform in mind. Items in the *device-dependent problems* category indicated issues beyond developers' ability and point towards issues directly related to the device such as ergonomics and/or device capabilities. Lastly, items that were included in the *game-specific problems* category were identified to be dependent on the game itself and did not indicate player experience differences between platforms. Moreover, the mobility module from the playability heuristics set was mostly associated with device related usability problems and not with the game design aspects. The results from the tests suggested that only the *MO1* and *MO5* items of the heuristic set was associated with game-specific features which again proved to be inefficient for analyzing player experience differences between platforms.

Common usability evaluation methods such as SUS and NPS were included in the tests and proved to be useful for evaluating usability related issues during the game. Usability test scores indicated that there was a difference of experience between platforms. SUS scores indicated a significant difference between platforms in terms of usability and NPS scores demonstrated that, unlike mobile players, PC platform players would recommend the game to their friends and colleagues. When SUS results were analyzed per item, PC players showed a tendency to play the game more often mainly because of the comfort, challenge and pacing factors. Results from the SUS tests also indicated that mobile platform players' feeling of confidence was linked to memorization while SUS

scores for PC players linked the same feeling with various heuristic items. Lastly, for PC players, the learnability of the system was found in direct relation with players' efficiency in understanding the terminology, onboarding, ease of use and story features of the game.

To conclude, covering most aspects of player experience, this study presents a new structure of evaluating platform differences utilizing heuristics. From stakeholders of the gaming industry to developers and designers, this method of analysis could be effective for understanding the needs and expectations of players. We believe that there is a need for a novel approach utilizing the heuristic set for evaluating player experience differences between platforms. It is understood that by adapting existing heuristics and iteratively improving them via user tests specifically for analyzing the platforms, would seem to be efficient for analyzing these differences in experience. However, it should be noted that the study was limited in terms of comparing only two gaming platforms with a small sample group. Therefore, the presented multi-modal approach should be tested on different gaming platforms with more players.

# 7. DISCUSSION

User-based testing has dominated the game design and research literature which describe playtests as the most preferred method for evaluating games (Rouse, 2001; Fullerton et al., 2004; Schell, 2008). Albeit, current researches have concluded that the adaptations of usability inspection methods can provide valuable feedback to developers and designers which can also offer more efficient and agile method of application (Molich, Kaasgaard, & Karyukin, 2004; Dumas & Loring, 2008; Korhonen, 2016).

Correspondingly, the PX researchers have proposed several methods for evaluating the playability of games to not just measure the effects of game elements on player experience and player attitudes for games, but also to provide a guide for designers and developers for creating better experiences as discussed in Chapter 2. Majority of these studies have focused on categorizing game elements while evaluating the experience as well as offering new perspectives for proposing both generalizable and domain specific methods. However, these proposed methods were seldom validated through comparisons with playtests or other measures such empirical results. As Wixon (2010) indicated, this lack of comparison affected the plausibility of the proposed methods. It was observed that none of the studies has managed to efficiently validate these propositions and every so often claimed that these methods are merely tools for evaluating some underlying playability aspects of the game within the broad scope of player experience.

In the reviewed literature, only a handful of different approaches have been implemented during player experience studies including; playtests, heuristic evaluations and literature reviews. However, a number of these researchers indicated that they would favor a multi-modal approach for validating the proposed methods (e.g. Desurvire, Caplan & Toth, 2004; Pinelle, Wong & Stach, 2008; Jegers, 2009; Hochleitner, 2015). On the other hand, only a few of the studies are observed to have conducted research on

evaluating player experience on different platforms (Suhonen & Vaataja, 2010; Gerling, 2011; Zaman et al. 2010; Rafaele, 2015; Kokil & Sanchez, 2015). With the increasing rivalry between different games, not only the intrinsic game elements matter but also the gaming platforms have started to play a crucial role. A number of researchers realizing this trend for cross-platform experiences have conducted studies for analyzing some of the differences between platforms; yet, all of these studies have focused on control mechanics when it comes to measurements as discussed in Chapter 3. Therefore, the lack of inclusive player experience evaluation methods for analyzing differences between platforms is the focus of this study and considered as one of the most crucial aspects for the betterment of the research field and the industry. Among various aspects of player experience, this study focuses on analysis between different gaming platforms.

### 7.1 Player Experience Evaluations

After extensive literature reviews and analysis, it was possible to categorize methodological differences between player experience evaluations. Four current modalities for evaluating games were identified; usability testing, empirical evaluations via questionnaires, playtesting and heuristic evaluations.

In terms of playtesting, the literature indicated that most of the evaluation studies conducted in the field of games have been utilizing playtests. They are based on fundamental notions of user testing and include similar procedures. However, it has been understood that playtests are effective for getting insights regarding playability problems and can focus on players' behavior as well as attitude towards the game. Hence, playtests are considered to be effective for measuring the goals towards achieving player experience goals in the literature (Davis et al. 2005, Fullerton, 2008). Most importantly, because of its potential for providing detailed feedbacks and observations form the players, playtests are often referred for validation purposes of other methods (e.g. Desurvire, 2004; Jegers, 2009; Korhonen, 2010). There is no commonly agreed timing for employing playtests for evaluating player experiences. Fullerton (2004) mentioned that playtests should be employed during the design process for better improving the game while Rouse (2001) argued that employing the playtests early could be obstructive for game development

process. Nevertheless, the literature reviews showed that a number of methods were employed for conducting playtests, including; observations, interviews, surveys, free-form evaluations, analysis of audio-visual recordings, physiological measurements, focus-groups, and utilization of game metrics. Clearly, not all the studies which used playtests applied all these procedures at once. It is observed that playtests most commonly include observations, interviews and audio-visual recordings. Therefore, these three playtest procedures were embedded into our tests.

In the first stage of this study, Playability Model (Sánchez et al. 2009) was employed for analyzing usability of games in an experimental test. This model was fundamentally adapted from conventional aspects which define usability; effectiveness, efficiency and satisfaction (ISO 9241-11: Guidance on Usability, 1998). However, the Playability Model does not claim to cover the specific aspects related with platform differences. In order to cover these platform-specific aspects, Mobility heuristics derived from Korhonen and Koivisto (2006) were additionally implemented during the tests. Furthermore, in order to gather quantitative data, a mobile eye-tracking device was used during the tests. This allowed to identify points of interest of the eye gaze during gameplay and to associate relevant feedbacks received from the players.

In the second stage of the study, it was identified that in recent years, questionnaires and scales have been developed to evaluate player experience in games (e.g. Takatalo 2007; Gamez et al. 2010; Gutwin et al. 2016; Depping et al. 2016). Poels, Ijsselsteijn & Kort (2008) proposed the Game Experience Questionnaire (GEQ), which offered a comprehensive and validated framework. It was identified that GEQ had an overall validity based on the semantics of the items. The questionnaire was tested by focus groups with a sizeable participant group and was described as providing reliable subscales for internal consistencies. Nevertheless, the questionnaire was not compared to playtests in terms of its effectiveness for evaluating player experience. To cover the approach of utilizing questionnaires for analyzing player experiences, a multi-modal approach utilizing GEQ along with playtests was employed.

Lastly, regarding the heuristic evaluations, it is understood that most of the studies for analyzing player experiences have utilized heuristic evaluations. This method of evaluation is considered to be beneficial since it requires less participants and provides application in a fast and cheap manner (Korhonen, 2010). Compared to playtests, heuristic evaluations neither require task oriented tests nor interviews. The literature review indicated that current heuristics did not offer a comprehensive and strongly validated set of items for evaluating differences between gaming platforms. In fact, previously conducted studies evaluating specific experiences proposed novel heuristics but only with the focus on a limited sense of analysis such as evaluating the game genre or the platform. These heuristics were developed to evaluate specific domains such as tabletop games (Köffel & Haller, 2008), mobile games (Korhonen & Koivisto, 2006) or educational games (e.g. Guo & Goh, 2016; Barbosa, Rego & Medeiros, 2015; Omar & Jaafar, 2010). Four different methodological heuristic approaches were found in the literature review; empirical evaluations, expert evaluations, literature reviews, and mixed-method approaches. The literature review indicated that multi-modal approaches such as combining playtests with heuristic evaluations yielded more and better detailed results in comparison to evaluation methods being employed individually. Korhonen (2010) underlined the importance of this issue by suggesting that several methods should be used as complementary to each other. At the end of the heuristic evaluation literature review, a list of 15 commonly employed heuristics by researchers were identified. When analyzed, Playability Heuristics (Korhonen, 2016) were found eligible for application of tests since it included these 15 common heuristic items as well as the modular structure for analyzing platform-specific elements. To cover this approach of utilizing heuristics for analyzing player experience, a multi-modal approach utilizing Playability Heuristics (Korhonen & Koivisto, 2006) along with playtests was employed.

Consequently, three variations of experimental studies were designed in order to assess player experience in different platforms inclusively; usability based evaluation, empirical evaluation and heuristic evaluation (Figure 7.1). All the tests were planned to follow a multi-modal framework, combining playtests with the suggested evaluation methods and aimed to propose an inclusive approach. During the tests, two casual games (*Fruit Ninja* for the first study and *Plants vs Zombies* for the latter two studies) were

examined. Casual games were chosen for the tests because of their simple, plain and commonly accepted properties (Kultima, 2009). They also do not require time-commitment and dedication as well as skillful players (Rafaele, 2015).

| | STUDY 1 | STUDY 2 | STUDY 3 |
|---|---|---|---|
| **Methods** | **Usability Based Evaluation + Playtesting** | **Empirical Evaluation via Questionnaires + Playtesting** | **Heuristic Evaluation + Playtesting** |
| **Techniques** | Proposed Framework combining *Playability Model & Mobility Heuristics*<br>Interviews<br>Observations<br>Audio-visual Recordings<br>Task-oriented Instructions<br>Mobile Eye-tracking | *GEQ*<br>Interviews<br>Observations<br>Audio-visual Recordings<br>SUS<br>NPS | *Playability Heuristics*<br>Interviews<br>Observations<br>Audio-visual Recordings<br>SUS<br>NPS |

**Figure 7.1** Method Modalities and Employed Techniques of the Experimental Studies

## 7.2 Experimental Studies

### 7.2.1 Usability Based Evaluations

In the first experimental study, an analysis between mobile and Xbox with Kinect gaming platforms were conducted. A combination of Playability Model heuristics (Sánchez et al. 2009, 2012) and Mobility heuristics (Korhonen & Koivisto, 2006) were employed for analyzing player experience between platforms. These methods were chosen since the playability model propose an approach aimed to characterize playability of games in a holistic manner. The model presents characteristics of playability by utilizing existing knowledge of usability and adapted the attributes of usability through implying different meanings to them in the context of games. Nevertheless, the model did not offer an approach for analyzing platform related attributes. Therefore, the Mobility heuristics were additionally employed during the tests since these heuristics were intended to focus on platform specific playability issues. The experiment included task oriented goals, play

sessions, player observations and user interviews. This qualitative study indicated that this combination of usability testing and analysis through proposed framework proved to be applicable for the analysis of the game but did not cover some of the important playability aspects such as differences in interaction modalities which the interviews revealed. The study essentially offered an approach based on usability aspects. As a result, only several of the proposed items such as "*effectiveness*", "*learnability*", and "*motivation"* were observed to identify playability issues regarding the difference between platforms. Despite including several evaluation methods and means for analyzing player experience differences, during the tests only the playtests along with playability model items of "effectiveness", "learnability" and "motivation" proved useful for analysis. Since the eye-tracking cannot provide the reasons why a person looks at a specific area in its nature, the utilization of the mobile eye-tracker only allowed for supportive data to some extent and did not provide correlation between player behavior in regards to platform differences. Similar in result, the mobility heuristics did not indicate clear difference between platforms since the heuristics only indicated hardware related differences, and did not provide additional insights related to playability differences regarding game elements.

The "effectiveness" item proved to be fulfilled well by the mobile platform since the players were observed to be more successful on the basis of their average completion times of the given tasks. The touch screen controls were identified to be superior when compared to full body gestures. Additionally, players mentioned that they felt more successful during mobile play sessions. While the Kinect peripheral offered a novel interaction modality, the learnability item was found satisfactory. This was mainly because of the intuitive nature of Kinect peripheral and the naturalness of the controls provided by it. However, players encountered playability problems regarding the controls during tests on console platform and felt frustrated because of them. As a result, the mobile platform was found to be less difficult to interact with compared to console platform. This indicated that the aspects of learnability and efficiency was useful for identifying control related playability problems in terms of platform differences.

Furthermore, in terms of "motivation", players were observed to be more motivated by the novel controls and social responses. The only observed reason for the players to lose motivation was the loss of control because of some interaction problems during the

game on console. Playing in front of friends increased the motivation for playing the game again, highlighting the importance of competitiveness in console platforms. On the other hand, mobile players were not observed to be competitive among themselves. These results indicated that the item of motivation might potentially be useful for identifying differences of experience between platforms.

It was assumed that the “immersion” item was predicted to be satisfied by the console platform since it offered a bigger screen size and an effortful but novel interaction type. Nevertheless, players did not felt immediately immersed to the game and were observed to be more stressed while playing the game on the console mainly because of the social pressure. A certain degree of correlation between social aspects, such as friends watching the player interact with the system, and immersion were observed. Because of this social influence affecting the dimension of the immersion, this item was not clearly observed for both of the platforms. Moreover, the non-isolated environment was observed to affect players and cause distractions. Hence it was not possible to analyze the item clearly. Furthermore, perhaps, the most constant aspect during the tests was social interaction. The item of satisfaction was identified in relation with the item of socialization. However, the socialization of players also had a negative impact on the analysis process since the behavior of the players differed when they are in a sociable environment.

During the analysis, it was identified that mobility heuristics included only a limited number of heuristic items regarding crucial differences such as the differences in screen sizes or peripherals. During the evaluations, mobility heuristics were able to indicate differences in terms of controls. These heuristics could be considered in association with efficiency and learnability.

In conclusion, only the playability model items of efficiency, learnability and motivation provided explicit implications for identifying playability differences between platforms. Additional items/heuristics have been found to be necessary for the identification of playability problems between platforms. Furthermore, observations and interviews have demonstrated to be crucial for the analysis of player experience. Most

importantly, this first study paved a way for further studies by presenting a basis for analysis of player experience between platforms.

### 7.2.2 Empirical Evaluations through Questionnaires

For the empirical evaluations utilizing questionnaires, a multi-modal research was designed, which included the GEQ questionnaires and playtests. PC and mobile gaming platforms were compared during the tests in terms of player experience. Essentially, the playtest methods were aimed to enrich the data gathered via GEQ questionnaire. In addition Nacke, Schild and Niesenhaus (2010) suggested that the questionnaire required the evaluation of usability of the game. Therefore conventional usability scales of SUS and NPS were additionally administered during the tests. Following the implications from the first experimental study and literature reviews, the playtests included semi-structured interviews, observations and audio-visual recordings. Derived from the lessons from the first experiment, the tests were conducted in an isolated environment to avoid any distractions and participants took the tests individually for eliminating the possibility of probable effects on social influence. Furthermore, unlike the first study, participants were recruited from university students but without prior experience with the game. This is because the prior knowledge of the game is considered to be potentially influential on players' attitude on reporting the experience (Poels, Kort & Ijsselsteijn, 2012). Moreover, Fullerton (2008) asserted that the players for the playtest sessions should be chosen from those who represented the target group of the game being tested. Because of this reason, only the players considered to be casual gamers were invited to the tests. All of the participants had prior experience in using smartphone technologies.

The comparison of GEQ scores between platforms did not provide significant summative difference in general. Only the SUS indicated a significant difference while GEQ questions were fundamentally subjective in structure and measure attitudinal preferences. However, central tendency of descriptive measures regarding GEQ revealed that the players on the PC platform reported more positive feelings such as feeling more successful, competent and immersed. It should be noted that GEQ questions were fundamentally subjective and were presented to measure players' attitudinal preferences.

Therefore GEQ should not be considered as a tool for summative results. The correlations among GEQ items indicated several important distinctions between platforms. The analysis showed that the items of flow, competence, and challenge were indicative components of player experience in different platforms. Furthermore, it was found that increase in challenge caused PC players to feel more negative feelings such as greater tension. On the contrary, mobile players were found to be motivated by the increase of challenge during the tests. Mobile players were also found to be more tolerant for coping with the demands of the game compared to PC players.

The interviews and observations proved to be essential for analysis of the GEQ results since they provided complimentary insights and suggestions regarding individual gaming experiences. Moreover, the results obtained from the interviews were categorized and grouped in relation to the GEQ items. This enabled the further analysis of the responses and indicated possible overlapping relations between interviews and GEQ scores. By the analysis of the interviews in relation with the descriptive results of the GEQ items, it was possible to identify main factors effecting the differences in terms of player experience between platforms as well as specific playability problems.

Moreover, further analysis showed that the usability scores did differ between the PC and mobile platforms and were compatible with some of the GEQ items. These scores were associated with the immersion and positive effect items on both platforms but on the mobile platform, these items were found to be significantly affecting the usability scores. Additionally, the usability scores were found positively related to the item of flow and negatively related to the items of negative affect and tension/annoyance for the mobile platform. This finding supported the notion that the flow on the mobile platform increases players' tolerance for usability problems.

In terms of methodology, three items of the GEQ (flow, competence, challenge) and the employment of SUS were found useful for analyzing player experience differences between platforms. Additionally, the playtests demonstrated to be influential in presenting an in-depth comprehension of different motivations of players on both platforms.

### 7.2.3 Heuristic Evaluation

The third study was based on the utilization of heuristic evaluation for analyzing player experiences. The aim of the study was to adopt a multi-modal approach to analyze differences of player experience between PC and mobile platforms. As suggested in the literature review, a multi-modal approach combining playtests with heuristics evaluations was followed. The analysis of the literature additionally indicated that Playability Heuristics (Korhonen, 2016) were qualified for application during the tests because of its inclusive structure covering the common heuristics in the literature and modular structure. It was also considered to be most valid (Paavilainen, 2010) due to its modular structure. The playability heuristics items were presented to be generalizable between platforms and game genres as well as mobility module specifically focusing on mobile games. The heuristic set was rated on a five-point Likert scale to make the evaluation process feasible for the evaluators. Detailed feedbacks regarding the heuristic items were gathered through semi-structured interviews.

As proven to be effective from previous experiments, semi-structured interviews, observations and audio-visual recordings were gathered in addition to the questionnaire and usability scales during the tests. Similar to the previous study, the tests were conducted in an isolated environment and participants were tested individually. The participants were recruited based on their gaming habits and their prior experience with games. Only the hardcore and mid-core players who have never played the evaluated game before were recruited for the tests. The reason for this was to prevent the possibility of prior experience of the game from affecting the player experiences in general and to simulate the expert review process.

Analysis indicated that the heuristic set was ineffective for analyzing different experiences between platforms when applied alone. The statistical analysis of the playability heuristic scores did not provide any significant difference between items.

However the comparisons of descriptive statistics with qualitative analysis utilizing playtests provided distinctive findings regarding platform differences. Thus, the player feedbacks proved to be invaluable for analyzing the experience differences.

Main contribution of the heuristic evaluation study was the identification of problem groups and their associations with the suggested playability heuristic items. As opposed to the common approach in which the identification of heuristics is often proposed through categorization of game elements, it was determined that this approach was insufficient for analyzing player experiences in different platforms. Similarly, Federoff (2002) mentioned that this model of categorization might not be directly related with the causes of playability problems. Supporting this notion, the feedbacks and observations suggested that each heuristic related to a specific playability problem that was mentioned through playtests. At first, the descriptions of problems during the tests were categorized. Similar to approaches proposed by Hara & Ovasaka (2014) and Soomro et al. (2012), the feedbacks describing the playability problems were grouped into three distinct categories: *Design-Dependent Problems*, *Game-Specific Problems* & *Device-Dependent Problems*. These categories were defined to provide insights about the issues that participants stated about the game and possible solutions to the problems encountered in the context of specific problem areas and were mapped with the heuristic items. Analyses for each problem group were conducted in detail revealing that several of the heuristic items were efficient for identifying different player experiences between platforms while the rest of the heuristic items were found to be non-eligible. Accordingly, four heuristics from the Game Usability module, nine heuristics from the Gameplay module and two heuristics from Mobility module were found to be eligible for evaluating different player experiences between platforms. It was indicated that the mobility module heuristics were not developed for evaluating experiences related to platform differences yet two of them proved to underline some playability issues that players encountered on mobile platform.

Furthermore, the usability scales of SUS (Brooke, 1996) and NPS (Reichheld, 2003) provided additional information. The results suggested that NPS survey helped identify the general tendency for the players to recommend the game in general meanwhile SUS tests indicated platform differences in terms of usability. To assess more details regarding

the scores of SUS, the SUS questions were analyzed individually. Three questions (1st, 9th and 10th) were identified as supportive to the observations.

In the end, a total number of 15 playability heuristics were found to be eligible for evaluating player experience differences between platforms. Additionally, the SUS proved to be instrumental for analyzing the general usability differences between platforms. During the study, the NPS scores indicated a difference between platforms yet in the previous study, the scores for both platforms were observed to be the same. This resulted in eliminating the scale as a reliable method for now until more experiments are conducted in the future.

The heuristic evaluations indicated that most of the playability issues were related to specific platform features such as the screen size and interaction modality, yet the source of these playability problems were found to be in other problem categories rather than *device-dependent problems* category. Furthermore, it was identified that majority of the playability problems addressed in *design-dependent problems* category could be solved via designing the game elements specifically for the targeted platform. These findings suggest that an additional platform based heuristics module should include features related to hardware of the device (such as, issues related to the ergonomics of the device, sound and output quality, screen size) as well as platform specific dynamics (such as issues related to control mechanisms or modalities of interaction) to inclusively evaluate the tested platforms.

### 7.3 Towards Constructing a Methodological Set

The studies conducted during the dissertation process revealed different factors effecting the player experience between platforms. Consequently, the iterative design of the studies allowed a comprehensive evaluation of the player experience. From the point of usability based evaluations utilizing playability model and mobility heuristics; “efficiency”, “learnability”, and “motivation” attributes were identified as the key factors for analyzing the player experience between platforms while the GEQ items of “flow”, “competence” and “challenge” were eligible for observing player attitudes regarding

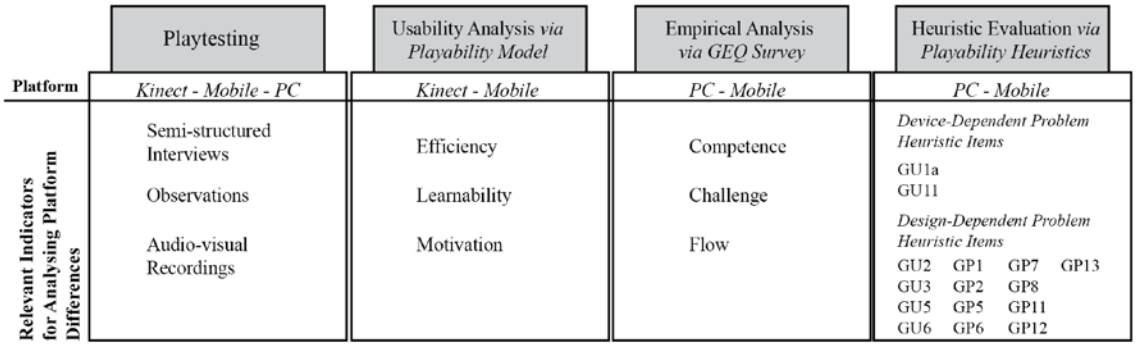



**Figure 7.3** Relevant Items Indicating Platform Differences

platform differences. Lastly, a selection of 15 heuristics from Playability Heuristics were found to be eligible for identifying more specific playability issues which indicate experience differences between platforms. Hence the combinations of these indicated items were considered as points of interest when analyzing player experiences between platforms (Figure 7.2). It was additionally found that the application of different methodological approaches combined along with playtests was an efficient method for analyzing the player experience in detail. Therefore, parallel to the findings in the literature (e.g. Papaloukas, 2009; Jegers, 2009; Desurvire & Wixon, 2013), this multi-modal approach should be employed during the evaluations for analyzing player experience between platforms.

In terms of utilizing these methods, two different implications are derived from the studies; implications on evaluation methods and on procedures. These implications are aimed to provide better evaluations for analyzing platform differences.

### 7.3.1 Methodological Implications

Wixon (2010) mentioned that, many of the presented approaches for evaluating user experience in games claim to offer similar advantages such as deeper insights or actionable data. Yet, indicated that their comparative usefulness was not clear. Moreover he mentioned that one of the most important factors that has not been emphasized by the field is the face validity or plausibility of the evaluations. Moreover, Duh et al. (2006) indicated that one critical factor in terms of validity of the evaluation is dependent on how

much the simulated scenario of the tests can present real life situation. This study aimed to address these problems through combining a number of methods and evaluations.

It has been noted a number of times in the literature that one cannot directly assess the usability of an artifact using a single measure (e.g. Nielsen, 1994, Ritter et al. 2014). Hence, the step-by-step utilization of multi-modal evaluation methodology employed in this study not only can provide cross-validation between the administered evaluation methods, but also enable analysis of playability evaluation techniques in a deductive manner. Furthermore, the multi-modal approach provides an opportunity for combining both formative (such as heuristics) and summative (such as usability scales) evaluations. Consequently, the multi-modal evaluations and the combination of different methods from different perspectives presented the grounds for proposing an inclusive set of evaluation methods for analyzing playability differences between platforms. Each study conducted during the preparation of the thesis involved different levels of detail in terms of player experience and playability. The findings of our research was compatible with what Bernhaupt (2010) mentioned as six distinct methods for evaluating player experience and playability of games is identified from the previous literature;

- Playtesting
- (Semi-structured) interviews
- Observations
- Audio-visual recordings
- Quantitative comparisons of gamers' behaviors via questionnaires focusing on users' attitudes, experiences.
- Heuristic evaluations (including heuristics for playability)

All of these methods were employed during the studies within three different experimental tests (Figure 7.1).

Face validity is assessed by having people such as experts or participants judge the proposed structure or test (Ritter et al. 2014). The combination of these methodologies and the tests conducted through the study provided face validity for the employed methods to the extent of their capability of measuring experience differences between platforms.

Additionally, since the methods employed in this study were derived from the literature, all of the methods had face validity because of their previous applications in the field of games. Additionally, our studies indicated that the results from the playtests were in correlation with the employed evaluation methods such as the GEQ questionnaire or Playability Heuristics. Although the methodologies employed during the study were found to be ineffective for analyzing platform differences when employed individually, the application of three different methodologies consecutively is found to be also ensuring internal validity, since each multi-modal evaluation experiment corresponded with each other to some extent. The literature review in this study provided a systematic approach covering the evaluation methods for analyzing player experience and playability problems. The chosen methods in the study for evaluating games and measuring different dimensions of playability and player experience indicated results on various dimensions of playability factors such as challenge or effectiveness. This enabled to inclusively evaluate different areas of player experience. Additionally, the relationship between playtests and other evaluation methods such as heuristics were observed to be directly relatable and comparable. Accordingly, playtests are found to deliver unstructured data which can be then structured through various evaluation methods in a multi-modal approach. This correspondence with the player feedbacks and playability evaluation items supported the internal validity of the overall methodology.

In the first experimental study, the utilization of the Playability Model (PM) allowed us to incorporate the proposed attributes of playability. Based on conventional usability attributes, the model is considered as a framework for analyzing playability in a broad sense. As Ibrahim et al. (2014) mentioned, it should be emphasized that playability model took the notion of playability as the usability in video games context (understanding and control of the game system), neglecting various elements affecting the experience in video games. Moreover the model was prepared for characterization of playability features and attributes (Sanchez et al. 2009; Sanchez, 2011). Thus, we additionally incorporated the “mobility heuristics” (Korhonen & Koivisto, 2006) which were administered during the analysis along with the playability model in order to analyze playability issues specific to platform differences. Through the experimental study, we found that only the attributes of “effectiveness”, “learnability” and “motivation” were the relevant attributes of playability

regarding platform differences and mobility heuristics indicated only hardware related issues. When the interview and observations from the study were inspected through the lens of these playability attributes, it was identified that the only the problems related to the controls of the platforms were included in the “efficiency” attribute. Moreover, the most dominant factor associated with the “learnability” attribute was also found to be in relation with the playability issues that are caused by controls. In the literature, the association between challenges in the game and “learnability” attribute is mentioned as, the challenges should be introduced in a way that give players the opportunity to study their behavior (Laramee, 2002). However, the playability problems regarding the controls were observed to disrupt this opportunity for the players to adapt to the challenges presented through the game, disrupting the learning process of the players. Lastly, the “motivation” attribute was identified to be in close relation with these two items, affected by playability problems related to controls and challenge. Players felt less motivated whenever they have encountered playability issues related to controls or challenges. It was found that the user preference of the platform relied upon the number of playability issues related with the controls and not the other attributes of playability such as immersion, emotion or satisfaction. Even though the players mentioned a bigger screen provided better immersiveness, they underlined that they would prefer to play the game on the smaller screen because of the less problematic controls. Thus, the study highlighted the importance of playability issues affecting the platform preference. Lastly, the playability issues related to the challenge of the game were found to be associated with the “learnability” attribute since the attribute was also defined to include the difficulty aspect of the game (Sanchez et al. 2009). The tests showed as players felt more control towards the platform, the more they found the game easy to play and felt less frustrated. Thus, the playability issues related to controls affected the challenge aspect. Although the study indicated that the motivation attribute resembles the platform differences to some extent, it did not provide additional findings rather than the “effectiveness” and “learnability” attributes does. In the end, from the findings of the first study, it is found that three of these defining attributes of playability differences were mostly related to the playability issues regarding the control of the system. Moreover, the interviews and observations played a critical role for associating the playability attributes and with player feedbacks. Additional findings indicated that the players were losing interest in the game because of

the playability problems during the game. Comparative results and feedbacks indicated that the platform preference was less associated with the attributes such as immersion and more associated with the number of playability problems encountered. It is found that the model offered a perspective for analyzing the platform differences in terms of challenge and control. Although the utilization of the model allowed for structuring playability problems in terms of playability attributes, it was not efficient for finding playability problems by itself. In addition, mobility heuristics identified hardware problems such as calibration problems with the Kinect peripheral. The results highlighted the association between the playability problems and the notion of balance between challenge and control. Hence it is found that the utilization of the employed methodology provided a general point of view towards inspecting differences between playability attributes on different gaming platforms, yet it was found not efficient for pinpointing specific playability problems such as issues related to interface or navigational mechanics of the game which were mentioned during the interviews.

In the second study, a more in-depth analysis regarding players' feelings towards the platforms were conducted. Comparisons between gamers' behaviors focusing on player preferences were analyzed using the GEQ questionnaire. Playtests including semi-structured interviews, observations and audio-visual recordings along with the SUS and NPS usability scales were also administered to procure and determine whether there are differences in terms of usability of the game between platforms. During the study, only the GEQ items of "flow", "competence", and "challenge" were found to be the marker components that differ due to the players' game experience on each platform. The "competence" item was related to the players' ability to control the game elements, while the "challenge" item referred to the balance in difficulty of the game. This indicated a similarity with the previous study in which the marker components for identifying platform differences were identified as control and challenge. When the results of the test were compared side-by-side with the previous findings, it was identified that the more immersed the players feel, the more susceptible they become to playability problems and less error-tolerant. Hence, it was discovered that as the immersiveness increased, the feeling of flow became more susceptible to disruptions from the playability problems. Furthermore, the more control the players feel, the more challenge they were looking for.

Thus, the findings underlined the importance of balance between challenge and control which fundamentally refer to “flow”. Usability scores also contributed to this finding. Considering that flow is defined as being dependent on the balance between frustration and boredom, the results showed that providing flow on the mobile platforms allows the players to tolerate playability issues. Therefore, it is found that the components of “flow”, “competence” and “challenge” from the GEQ questionnaire offered a different and broad perspective for analyzing the difference in player experience between platforms in terms of players’ feelings and attitudes. In relation to previous study, the findings indicated parallel notions regarding the playability differences between platforms yet the utilization of the questionnaire allowed for analysis of players’ feelings. The employment of SUS usability scale (Brooke, 1996) additionally provided summative and supportive results regarding the platform differences. Although the comparisons of GEQ scores between platforms did not indicate any significant difference, the SUS scores have. Thus, it can be speculated that the SUS was efficient for identifying whether there is a difference between platforms in terms of playability. Although the multi-modal approach and empirical evaluations along with playtests provided more inclusive and detailed insights regarding the player experience differences between platforms, the GEQ questionnaire lacked the ability for identifying specific playability problems in detail and yielded no results for identifying possible solutions to those problems. As a result, it is found that the utilization of the questionnaire provided a point of view towards inspecting players’ attitudes and attitudinal differences between platforms while playtests with semi-structured interviews, observations and audio-visual recordings were instrumental for identifying the causes of these differences.

In the third phase of the study, derived from previous findings, a heuristic evaluation study utilizing Playability Heuristics (Korhonen, 2006) was designed with the notion of identifying different playability problems between platforms. Playability heuristics were developed for assessing playability problems of the games focusing on mobile platforms. It should be noted that the aim of this study was not the identification of general playability problems of the game but to indicate platform related differences. Similar to the previous study, playtests including semi-structured interviews, observations and audio-visual recordings were additionally employed during the tests. The results indicated several

implications on playability issues regarding the game but only some of them were found to be in relation to platform differences. Additionally, unlike previous research, utilizing heuristics proved to be effective for indicating a number of specific playability problems when taken in consideration with playtest results. However playability heuristics were found to be inefficient for analyzing platform differences when applied by itself. Nonetheless, the heuristic set provided a guideline for analyzing and structuring different game elements through playtests such as revealing comparable playability issues regarding the interface design or interaction design. The modular structure of the heuristic set was found to be inefficient for analyzing differences between platforms since many of heuristic items related to a specific module were in fact found to be in association with other relevant modules. We identified that players do not relate heuristic items directly with their given modules such as a number of “gameplay” module items referring to usability problems instead. Nevertheless the descriptive measures of the heuristic set and the qualitative analysis results of the playtests were associated during the study, thus provided the possibility to categorize three distinct problem classifications, and various playability issues under these categories (Device-dependent, Design-dependent & Game-specific problems) Previous research indicated the importance of flow, challenge and control aspects in general yet they did not indicate any specific playability issues. Moreover previous tests were not capable of identifying specific playability problems related to platform differences. The heuristic items along with the proposed categorizations were not only found useful for classifying specific playability problems but also provided a broad perspective for looking into the associations between causes and potential solutions of the playability issues in relation to platform and game features. Thus, categorization of playability problems have potentially increased the applicability of the method being used in the gaming industry.

In sum, this study and the employed methodologies through experimental studies was found to be effective for identifying the underlying playability problems regarding the differences between platforms. Thus, the triangulation of methods was useful for identifying the reasons of the player experience differences between platforms inclusively and cross verification of the findings. Different methodologies were discovered to identify different factors regarding the player experience. Playability model was found to focus on

playability attributes while the GEQ questionnaire was more focused on players' general attitudes and feelings. Lastly, the playability heuristics were effective for evaluating specific playability problems. Moreover, the findings from the experimental studies provided an internal validity to some extent since cause and effect relations were identified through the study utilizing a number of different methodologies while the study conditions were regulated in laboratory conditions as advised in the literature (Ritter et al. 2014).

As a contribution, the analyses utilized through the study from a deductive point of view could prove to be instrumental in game development and evaluation processes focusing on platform differences, for example by the developers choosing the best evaluation methods and techniques based on the desired detail and precision of playability problems.

**Table 7.4** Differences between Methods and Scopes of Analyses

| | **Generic** ⟷ | | | **Specific** |
|---|---|---|---|---|
| **Evaluation Type** | Summative | Formative | Formative | Formative |
| **Goals And Scope of Analyses** | **Determining** player experience difference between platforms | Analysis of **player attitudes** and **feelings** between platforms | Analysis of **playability attributes** between platforms | Analysis of **playability** between platforms and **categorization** of problems. |
| **Method(s)** | System Usability Scale (SUS) | Game Experience Questionnaire (GEQ) + Playtests | Playability Model (PM) + Playtests | Playability Heuristics + Playtests |

According to the findings from these studies, it is found that different methods offer different analysis scopes for analyzing player experience between platforms (Table 7.1). In general, the SUS usability scale is found to be useful for identifying whether there are playability differences between platforms in a summative fashion. Moreover, the

playability model employed in the first experiment is found to be useful for providing a broad sense of differences between platforms in terms of playability attributes based on the usability of the game. Following this, the GEQ questionnaire is found to be effective for analyzing player attitudes. The results of the second experiment could be associated with the previous findings, presenting a better perspective into the previously indicated playability attribute differences through the lens of players' feelings. However, combination of these two methodologies necessitate the utilization of playtests for the validation and in-depth analysis of the results since the questionnaire only allows to gather quantitative scores and lack qualitative data. Lastly, the third stage of the study showed that playability heuristics were useful for analyzing specific playability problems between different platforms for which the previous studies were not capable of. Thus, the employment of heuristics provided a crucial insight through differing playability problems between platforms. However it should be noted that playability heuristics neglected the players' feelings and attitudinal preference during the evaluations and should be considered as a tool for analysis of games in terms of playability only. The utilization of playtests including semi-structured interviews, observations and audio-visual recordings allowed for associations to be made between heuristics and player feedbacks, providing causes and potential solutions to those playability problems. Most of the research regarding the heuristics were found to be based on heuristics which are fundamentally based on game design and development areas such as mechanics, interface or gameplay (Federoff, 2002). Contrastingly, the categorization of the playability problems through playtest results enabled a separation of the identified playability problems from a functional perspective. The categorization of design-dependent problems, game-specific problems & device-dependent problems offer an applicable perspective to game developers and researchers since these categories indicate playability problems in relation to their dependencies and probable corresponding solutions.

### 7.3.2 Procedural Implications

The tests conducted during the study indicated that some of the procedures were found to be ineffective and/or obstructive for analyzing player experience differences and

playability problems between platforms. However the tests also highlighted the importance of a number of the procedures that have been utilized during the study.

In the first experimental study, pre-tests were conducted for assessing demographic information of the participants which yielded no observable relation with the test results. Hence it was found to be ineffective during the first stage of the study. Similarly, the application of physiological measurement via eye-tracker was found useful for associating feedbacks received from the players but did not provide indicative results by itself. Previous researches on player experience (e.g. Ijsselsteijn et al. 2012) suggested that player experience may be influenced during the gameplay sessions by factors such as player types and gaming frequencies. It was mentioned that previous experiences can potentially be influential in player attitudes and skills, affecting the overall goal of the tests. Yet, during the pre-tests, previous experience with the game and gaming frequencies were not specifically addressed. In terms of gameplay, a task-oriented test was designed during the first experiment. This was observed to prevent the participants' willingness for exploration of the system and limit the overall game experience to some extent. Moreover, this prevented the analysis of the "discovery" attribute of the playability model. It was also found that the non-isolated testing environment affected the player experience greatly. Besides, it was identified that the social pressure and/or the feeling of being watched by friends affected the participants causing them to behave differently. Players were allowed to watch and talk to each other during the test sessions to allow analysis on the socialization item. Albeit, these circumstances affected the efficiency of analysis as a whole. As a result, it was understood that the evaluation of the socialization attribute with the rest of the playability attributes in the same environment was not possible and that the experiments would be best conducted in an isolated environment to prevent potential distractions.

Following the lessons learned from the first stage of the study, pre-tests were iteratively improved in the second study. Through them, player attitudes and previous experiences were identified. Therefore, only the players that have never played the game in question were recruited for the tests, eliminating the chance of players being familiar with the game before. Additionally, the participants were tested in an isolated environment

and tested individually to prevent any environmental distraction including social influence. Unlike the previous test, the players were not allowed to think-aloud since it was found to be affecting the immersion during the play sessions. It was understood that the playtests were much more effective for indicating specific playability problems compared to the survey method which was useful for indicating players' general attitude towards the game. Hence the multi-modal approach employed during the tests was observed to be effective for analyzing the player experience from both dimensions. The semi-structured interviews and audio-visual recordings demonstrated to be an efficient method for associating analysis results with the observations and the questionnaire. These playtesting methodologies were chosen according to their common application in the literature which was reviewed in Chapter 2. When compared with the first experimental study. The addition of semi-structured interviews enabled the possibility for forming associations with the questionnaire items and results.

Lastly, in the third experimental study, iterative improvements of procedures from previous researches were taken into consideration while designing the test. The isolated environment prevented any unwanted distractions influencing the player experience during the gameplay sessions. Similarly, think-aloud method was not employed during the research for measuring the immersion aspect of the game. Differently, instead of recruiting experts for evaluations as previous researches utilizing heuristics did (e.g. Desurvire et al. 2004, Ponnada & Kannan, 2012), hardcore and mid-core gamers were invited to the tests. This purposive sampling prevented the influence of previous experience with the game as well as allowing to conduct playtests and heuristic evaluations with the same participants which allowed the possibility to associate playtests feedbacks with heuristic results individually. Most importantly, the categorization of the playability problems and heuristics through player feedbacks in the analysis procedure was found to be invaluable for presenting a plausible perspective to game designers and developers.

Ultimately, it was found that a controlled environment was important for evaluating player experience of the games between platforms. Although a more natural and familiar gaming environment would support a simulated ambiance for the tests and enable

ecological validity, it is observed that the environmental distractions such as social influences and/or other related factors have an impact on player behavior and feedbacks. Moreover, it was understood that the pre-tests and purposive sampling was crucial for analysis since previous experience of games affect the general player experience of the player. Among various playtest methods, semi-structured interviews, observations and audio-visual recordings were found to be most efficient for analyzing player experience between platforms. Moreover, think-aloud research method was found to be distractive in nature, disrupting players' immersion. It is understood that the playtests were useful for gathering information regarding the gaming system in an unstructured order and were instrumental for identifying relations between playability aspects regarding the game and the platform.

### 7.4 A Proposition towards the Implementation of the Methodological Set

It should be noted that the games tested through this study were casual games. Games from this genre were selected since they did not require any time commitment or skill from the players. Moreover the games '*Fruit Ninja*' and '*Plants vs Zombies*' were found eligible for analyzing platform differences since these two games were developed for multiple platforms. Thus, the implications of this study should be considered viable for multi-platform causal games.

During this study, it was demonstrated that the SUS usability scale was efficient for determining whether there are different player experiences between platforms. It was found that different platforms offer different and various experiences even though the players interact with the same game. Moreover, these differences between platforms can be analyzed through proposed multi-modal methodologies, each underlining different factors of player experience. For an inclusive approach it was understood that the combination of usability based evaluations, empirical evaluations via questionnaires, heuristics evaluations and playtests were necessary for evaluating various factors affecting the player experience.

Fullerton (2008) argues that one of the essential processes during the development cycle of a game is the playtests. He indicates that there is a misconception about playtests which they are often considered as just playing the game and gathering feedbacks. Yet in reality, the play sessions are only a part of the playtest process. Moreover he underlines the importance of the application of playtests during the development process rather than after the product release since only then the playtest results would prove useful for iteratively improving the games. In summary, he indicated that even the prototype of the game should include iterative evaluation processes in order to deliver a high quality game. Nevertheless, he additionally mentions that the rather expensive cost for playtesting the game for many times during the design is considered to be unfeasible among game developers. In this study, one of the findings is that playtests deliver unstructured data which can be structured through various evaluation methods in a multi-modal approach. Moreover, this multi-modal approaches showed that only by examining some of the items from the proposed methods would suffice for analysis between platforms. Thus, the presented methods and their identified sub items in this study are demonstrated to be useful for analyzing certain aspects of the game during the development process without the need for an all-round and costly playtests in every developmental phase.

Fullerton (2014) divided the consecutive phases of game production into five levels; '*concept*', '*pre-production*', '*production*', '*quality assurance (QA)*', and '*maintenance*'. To provide a proposition regarding the implementation of the presented methodological set, the production stages of a game, which were indicated by Fullerton (2014) were associated with the methodologies and evaluation procedures for further guiding the researchers and developers in the industry (Figure 7.3).

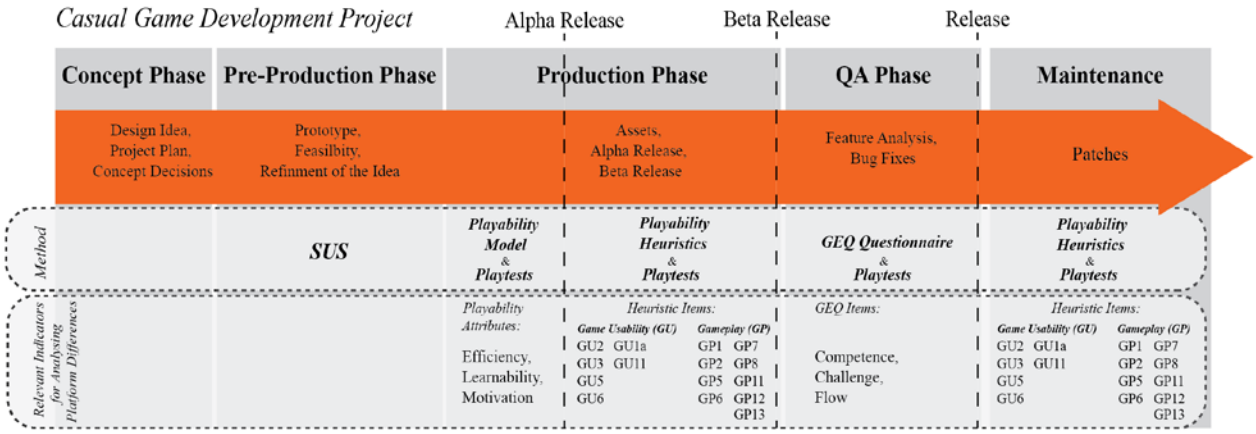



**Figure 7.3** Implementation Framework for Analyzing Platform Differences through a Casual Game Project

### 7.4.1 Concept Phase

The concept phase refers to the phase in which the developers have the most freedom for deciding on main concepts of the game. In this stage the developer(s) could propose major solutions and perspectives for developing the game. Fullerton (2014) mentioned that the major issues which would affect the overall of the game could be found in the concept and preproduction stages through evaluations. Moreover he mentioned that the concept stage of the development is crucial for pitching the game ideas to publishers. The subsections related to the game project, such as the development team, project plan or targeted platform(s) should be defined at this stage.

### 7.4.2 Pre-Production Phase

Following the concept stage, the pre-production phase includes the prototype design of the game, often testing the feasibility of the idea. Following this notion, Fullerton (2014) asserts that the refinement of the design and development of the prototypes in this phase would allow the developers to assess the feasibility of implementing "risky technology" elements. He underlines that this would help to "*reduce the potential risk for both developer and publisher*". Hence for multi-platform games, the targeted platforms should be tested during this stage. It was found that in the pre-production phase, by

employing SUS test after gameplay sessions with the prototype game, the designers could determine if there are player experience differences between the platforms and plan the development process accordingly in a cheap and fast manner. As a result it is sought that they would have the possibility to address whether the game should be designed considering separate player expectancies and capabilities of the platforms in terms of player experience or the same design would work between platforms. However, it should be noted that observation based qualitative techniques such as the playtests are recommended in addition to SUS surveys since they can provide insights and information that cannot be identified via the survey by itself.

### 7.4.3 Production Phase

The production phase is considered to be the longest phase of the project which includes the preparation of the game elements such as the code, sound design, arts and animations of the game. The alpha and beta versions of the game were released in this phase, delivering a playable game with all the major functions (Fullerton, 2014). Alpha and beta versions of the games refer to functional differences as in alpha release the game is expected to have majority of the functions working while in beta version, all the functions of the game should have been implemented. Beta version begins when all the essential game features were implemented. Before the alpha release, the development team has the possibility to identify and solve issues regarding the playability of the game. Therefore employing the playability model for evaluating the usability of the game before alpha release should be considered for indicating major differences between platforms in terms of playability attributes in a broad sense. For indicating platform related playability attribute differences, it is found that the evaluators should focus on the “efficiency”, “learnability” and “motivation” attributes of the playability model. This would allow developers to take action when or if a difference of playability is identified between gaming platforms with ease and without necessarily falling behind in the development schedule. Additionally, playtests are found to be necessary in this phase for identifying the different playability attributes between platforms in general. After the alpha release, the beta version of the game is often developed which generally refers to the feature-complete-game state but likely to contain known or unknown problems. Thus, before the

beta release the heuristic evaluations should be employed in order to identify specific playability problems between platforms. In order to conduct the heuristic evaluations, it is found that the 15 playability heuristic (Korhonen, 2016) items that are identified in the previous section, should be evaluated for analyzing specific playability differences between platforms. Although the heuristic evaluation methods are often considered to be conducted through experts in the literature, the study demonstrated that recruiting hardcore or mid-core gamers with game design experience to some extent is proved to be effective. Additionally, through utilizing the presented categorizations of playability problem dependencies, designers and developers would potentially address the playability problems through the perspective of these categories in a more direct, efficient, and solution-oriented manner.

### 7.4.4 Quality Assurance (QA) Phase

After the beta release of the game, during the QA phase, the focus shifts to the analysis of the functions as to assess whether they work in an expected manner and the artwork and levels are complete and polished. During this phase, the game is prepared for final release. In this stage of the development, the QA team inspects the game for any remaining bugs and deficiencies often conducting user-tests. Thus, during this stage it was found that the empirical evaluations using the GEQ questionnaire for assessing players experience differences between platforms should be employed. It was found that through analyzing the GEQ items of “competence”, “challenge”, and “flow” it is possible to indicate platform differences in the context of players’ attitudes and feelings. Additionally, by implementing playtests along with the questionnaire, it would be possible to re-evaluate the progress between previously conducted playability model evaluations and asses the improvements on playability issues identified in the production phase. Employing playtests along with the questionnaire would also be helpful to indicate associations with the GEQ items, providing an in-depth analysis of the player experience.

### 7.4.5 Maintenance

Lastly, after the release of the game, Fullerton (2014) mentions a maintenance phase in which the released game is often “patched”, generally fixing feature problems that manage to make it past the QA phase. These patches provide the possibility for solving specific issues related to platform differences via requiring small changes in the game. Thus, playability evaluations utilizing 15 playability heuristics that are indicated in the study should be employed before patch releases for identifying playability problems related to platform differences.

# 8. CONCLUSION

From simple arcade machines in alleyways to high end PC's at homes, games became ubiquitous in our lives. The exponential success and popularity of games have received well deserved attention not just from the industry but also from user experience researchers. Similar to the increasing importance of user experience and usability in productivity software, stakeholders and researchers have adopted the term player experience and playability for developing better games and for delivering better experiences. Hence following the conventional user-oriented approaches, researchers from various fields have started to study methods for evaluating player experience and playability of games. Since the games are designed in accordance with the new technologies and interaction capabilities in mind, gaming platforms became as important as the game itself. Today, most of the gaming companies compete for gaining place in the industry and they aim to achieve that by releasing games for multiple platforms. As it can be seen, platform differences between games play a crucial role in terms of player experience. However, there is still a lack of understanding on how the interactive capabilities of games relate to user experience and usability, and the player experience differences between platforms were not analyzed thoroughly.

In this study, one of the main research questions was whether there are significant differences in the gaming experience between platforms and if those differences can be explained by various descriptive methods covering the player experience for games. To answer that, a comprehensive literature review was conducted during the study, inspecting various playability evaluation methods and techniques for evaluating player experience in games. It is apparent that there are several fundamentally different approaches for evaluating player experience and playability of games in the literature. However it is also understood that to analyze player experience in detail, it is crucial to evaluate playability differences between platforms. For evaluating those diverse aspects of player experience

and playability, three separate but consecutive tests were conducted and potential frameworks were identified. These tests suggested significant variances between platforms in terms of presenting different viewpoints of analysis for each method administered. However, employing a particular evaluation method individually yielded no definitive results in terms of comparing player experiences between platforms and therefore necessitate other evaluation methods such as playtests, to be included during tests for extensive analysis. It is indicated that the majority of the player experience differences between platforms emerge from game design elements and not from the different hardware capabilities. The studies showed that designing and conducting a play sessions followed by employing the SUS test would determine whether there is a difference in player experience between platforms. The SUS is fundamentally a summative evaluation method and therefore could not provide formative evaluation insights regarding the identification or the solution of playability problems. Hence, the SUS tests were identified as eligible for only determining if there are any player experience differences between platforms by itself without requiring additional evaluation methods yet did not provide any details regarding playability issues and specific differences between platforms. Moreover, experimental studies indicated that for assessing playability differences between platforms in a broad sense, playability attributes of “effectiveness”, “learnability” and “motivation” from the playability model should be analyzed in addition to playtests. Interviews and observations during and after the play sessions provided the possibility of correlating playability attributes with the player feedback. Through application of the playability model and the playtests, it is found that the playability model determined playability differences related to controls of the game system and the challenge of the game. However, the model only offered a perspective towards analyzing differences in the given playability attributes and lacked the capability of pinpointing platform specific playability differences. For the analysis of these specific playability differences between platforms, the application of the 15 identified playability heuristics in the analysis was found to be effective. Along with heuristics, the playtests provided the chance to associate player feedbacks with the heuristics, enabling the possibility for categorizing problem relevancies according to their dependencies such as device-dependent, design-dependent and game-specific problems. This increased the applicability of the method since these categories indicate playability problems in relation

to their dependencies and probable corresponding solutions. Lastly, the application of GEQ questionnaire along with playtests enabled the possibility of evaluating differences between platforms in terms of player experience. However, similar to previously employed methods of playability model and playability heuristics, GEQ was found to be insufficient for analyzing player experience differences between platforms when applied by itself. Nevertheless, the statistical analysis and combination of GEQ tests with the player feedbacks indicated that the GEQ items of “competence”, “challenge” and “flow” were the marker components for player experience differences between platforms. It was indicated that these items of the questionnaire were related to the playability attributes of “effectiveness” and “learnability”, thus provided a relation between these two study results as discussed in the previous section. Moreover the questionnaire was found effective for evaluating player experience differences in terms of players’ attitudes and feelings, rather than differences of playability between platforms. In conclusion the combination of the evaluation methodologies along with playtests were found to be effective for analyzing different aspects of the player experience and playability differences between platforms.

The second research question of this study was whether it is possible to identify the differences in player experience between platforms inclusively by a proposed original multimodal evaluation. It is understood that the differences between platforms can be analyzed through the proposed multi-modal methodologies, each underlining different factors of player experience. In order to provide an effective and holistic approach for evaluating player experience between different gaming platforms, a selection of a number of distinct approaches including playtests, heuristic evaluations, quantitative player experience surveys, and usability scales were employed during the study. These methods were selected according to their relevance in the literature as well as their inclusive context for which the combination of these methods covered all the aspects mentioned in the literature as discussed in the previous section. This multi-modal combination of methods, individually or en masse, proved to be effective for structuring associations between gathered data, exploring the validity of the results, and thus providing a solid foundation for delivering effective analyses. We believe that the conceptual application of the

framework for proposed methodologies and their sub items, could potentially be helpful for both the game developers and researchers.

As games become more sophisticated and players were engaged and immersed in the game world, traditional usability evaluation methods seem to be inefficient for analyzing player experience in an optimal way. In this study, we indicated that through the application of these methods from a multi-modal perspective, the difference in player experience between platforms can be understood in a better way. It was seen that the research field offered many different approaches towards analyzing the player experience, often suggesting novel procedures and models, creating a clutter of methods. Accordingly, it was recognized that one of the main problematic areas of the research field was the plausibility and the applicability of the evaluation methods. Through forming associations between evaluation methods and player feedbacks this study covered this problem in terms of identifying relevant playability issues and their probable solution in terms of platform differences. Methodological triangulation in the study facilitated validation of findings through cross verification. Additionally, the study have shown that certain procedures can be obstructive or ineffective while a number of them are found to be crucial for assessing information on platform differences. It was understood that a controlled environment was necessary for effectively evaluating player experiences regarding the platform differences. Additionally it was found that a purposive sampling along with pre-tests was crucial for analysis because previous experience of games may potentially affect the general player experience. Lastly, it was found that playtests utilizing observations, semi-structured interviews and audio-visual recordings are useful for gathering unstructured information regarding the game system and were instrumental for identifying relations between playability aspects regarding the game and the platform. More so, by proposing an implementation framework of the studied methods, we aimed to present an applicable methodological set for stakeholders and researchers. Applicability of the proposed methodological set was inspected through production phases of a casual game development process. Ultimately, these production stages of a causal game were associated with the methodologies and evaluation procedures for further guiding researchers and developers. We believe that such a set potentially improve the field of

game research as well as gaming industry which in turn developers will provide better games with less playability problems because of platform differences.

### 8.1 Limitations of the Study

This study has some limitations such as the multi-modal test design and relevant procedures applied should be repeated with higher number of participants for further validation. Additionally, the proposed framework of methodologies was only tested on a specific game genre (casual game) and on limited number of platforms. According to the study, the playability problems between platforms vary depending on the comparison of platform preferences. Thus, additional platforms should be tested in order to validate the proposed approach.

Moreover, the chosen sampling method was purposive sampling, where the demographics, player skills and player preferences are specified. In purposive sampling, subjects are selected based on the purpose of the study. Although the participants were recruited from the university students that resemble the target audience of a causal game for the tests, they may not strictly be used to make inferences about wider populations. Therefore, the presented multi-modal approaches should be tested on different gaming platforms and with a larger sample who have different levels of gaming experiences. More so, the gender preference or a need for balance between genders was not specified for the study. As a result, the majority of the players that attended the tests were male participants. Thus, a more gender-based study might be more appropriate for understanding if there are any gender related variations in play across different platforms.

Fundamentally, this study was focused on evaluating casual games since they offer similar achievement for good results between platforms (Raffaele et al. 2015), and they do not require time commitment or skillful players. There is no precise classification of casual genres in the modern gaming industry. However, we indicated that different mechanics in different casual games may emphasize different focal points in terms of player and platform capabilities. Although the two tested games, *Fruit Ninja* and *Plants vs Zombies*, were indicated to be in the same casual game genre of “action & arcade” as

defined by Big Fish Games, one of the leading casual game developers and distributors, it was observed that the latter game, *Plant vs Zombies,* offered a number of game mechanics compared to the single-mindedness of the *Fruit Ninja* game. This may have affected the results since in the first phase of the study, which focused on the *Fruit Ninja* game, playability attributes were mainly related with the controls of the game while latter studies indicated different aspects affecting the experience as well. Hence, additional casual games focusing on different interactive capabilities may potentially increase the validity of the proposed methodology.

### 8.2 Future Work

This study is aimed to serve as a stepping-stone for evaluating platform differences in terms of player experience. The results of the study can be used as a guideline for game researchers and game developers focusing on evaluating cross-platform player experience. The results of the study indicated that the proposed methodological set presented in this study can be improved by the implementation of other evaluation methods covering various unmentioned aspects of the player experience such as player psychology or other external factors affecting the player experience such as social influence.

Furthermore, these findings can be used to develop a set of heuristics for evaluating cross-platform player experiences which can be utilized by game researchers and developers. My long-term goal is to develop such heuristics and provide a comprehensive set of heuristics using these results for analyzing platform differences through expert evaluations. I believe that such a set potentially will improve the field of game research as well as gaming industry which in turn developers will provide better games with less playability problems that are caused by platform differences.

## REFERENCES


Admiraal, W., Huizenga, J., Akkerman, S., & Ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185-1194.

Aker, Ç., Rızvanoğlu, K. & İnal, Y. (2017). A Multi-Modal Approach for Evaluating Player Experience on Different Gaming Platforms.

Aker, Ç., Rızvanoğlu, K., İnal, Y., & Yılmaz, A. S. (2016, July). Analyzing Playability in Multi-platform Games: A Case Study of the Fruit Ninja Game. In *International Conference of Design, User Experience, and Usability* (pp. 229-239). Springer, Cham.

Al Mahmud, A., Mubin, O., Shahid, S., & Martens, J. B. (2008, October). Designing and evaluating the tabletop game experience for senior citizens. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (pp. 403-406). ACM.

Almeida, S., Mealha, Ó., & Veloso, A. (2011). Heuristic Evaluation of "FarmVille". *PRISMA. COM*, (14), 44-58.

Baauw, E., Bekker, M. M., & Barendregt, W. (2005, September). A structured expert evaluation method for the evaluation of children's computer games. In *IFIP Conference on Human-Computer Interaction* (pp. 457-469). Springer, Berlin, Heidelberg.

Barkhuus, L., & Rode, J. A. (2007, April). From mice to men-24 years of evaluation in CHI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1-16).

Bernhaupt, R. (Ed.). (2010). *Evaluating user experience in games: Concepts and methods*. Springer Science & Business Media.

Bernhaupt, R., Eckschlager, M., & Tscheligi, M. (2007, June). Methods for evaluating games: how to measure usability and user experience in games? In Proceedings of the international conference on Advances in computer entertainment technology (pp. 309-310). ACM.

Bernhaupt, R., Ijsselsteijn, W., Mueller, F. F., Tscheligi, M., & Wixon, D. (2008, April). Evaluating user experiences in games. In CHI'08 extended abstracts on Human factors in computing systems (pp. 3905-3908). ACM.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*(4), 624-634.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Calvillo-Gámez, E. H., Cairns, P., & Cox, A. L. (2010). Assessing the core elements of the gaming experience. In *evaluating user experience in games* (pp. 47-71). Springer London.

Carmody, K. W. (2012). *Exploring serious game design heuristics: a delphi study* (Doctoral dissertation, Northeastern University).

Caroux, L., & Isbister, K. (2016). Influence of head-up displays' characteristics on user experience in video games. *International Journal of Human-Computer Studies*, *87*, 65-79.

Clanton, C. (1998, April). An interpreted demonstration of computer game design. In *CHI 98 conference summary on Human factors in computing systems* (pp. 1-2). ACM. New York, NY, USA

Clanton, C. (2000). Lessons from game design. *Information Appliances and beyond: interaction design for consumer products*, 299.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal performance*. NY: Cambridge University Press, 40.

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

Depping, A. E., Mandryk, R. L., Li, C., Gutwin, C., & Vicencio-Moreira, R. (2016, May). How disclosing skill assistance affects play experience in a multiplayer first-person shooter game. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3462-3472). ACM.

Desurvire, H., & Wiberg, C. (2009, July). Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration. In *International Conference on Online Communities and Social Computing* (pp. 557-566). Springer, Berlin, Heidelberg.

Desurvire, H., & Wiberg, C. (2015). User Experience Design for Inexperienced Gamers: GAP—Game Approachability Principles. In *Game User Experience Evaluation* (pp. 169-186). Springer, Cham.

Desurvire, H., & Wixon, D. (2013, April). Game principles: choice, change & creativity: making better games. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1065-1070). ACM.

Desurvire, H., Caplan, M., & Toth, J. A. (2004, April). Using heuristics to evaluate the playability of games. In *CHI'04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1509-1512). ACM.

Detenber, B. H., & Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. *Journal of Communication*, *46*(3), 66-84.

Dix, A., Finlay, J., Abowd, G. & Beale, R. (1997). Human-Computer Interaction (Second Edition). New York: Prentice Hall.

Duh, H. B. L., Tan, G. C., & Chen, V. H. H. (2006, September). Usability evaluation for mobile device: a comparison of laboratory and field tests. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 181-186). ACM.

Egenfeldt-Nielsen, S., Smith, J. H., & Tosca, S. P. (2013). *Understanding video games: The essential introduction*. Routledge.

Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analyzing immersion. *Worlds in play: International perspectives on digital games research*, *37*(2), 37-53.

Esposito, N. (2005). Game atmosphere archiving thanks to virtual reality for the preservation of the video game cultural heritage. In *Proceedings of ICHIM* (Vol. 5).

Fabricatore, C., Nussbaum, M., & Rosas, R. (2002). Playability in action videogames: A qualitative design model. *Human-Computer Interaction*, *17*(4), 311-368.

Febretti, A., & Garzotto, F. (2009, April). Usability, playability, and long-term engagement in computer games. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems* (pp. 4063-4068). ACM.

Federoff, M. A. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games* (Doctoral dissertation, Indiana University).

Fierley, R., & Engl, S. (2010, September). User experience methods and games: lessons learned. In *Proceedings of the 24th BCS Interaction Specialist Group Conference* (pp. 204-210). British Computer Society.

Fritsch, T., Voigt, B., & Schiller, J. (2008). Evaluation of input options on mobile gaming devices.

Fullerton, T. (2008). *Game design workshop: a playcentric approach to creating innovative games*. CRC press.

Fullerton, T., Swain, C., & Hoffman, S. (2004). *Game design workshop: Designing, prototyping, & playtesting games*. CRC Press.

Garzotto, F. (2007, June). Investigating the educational effectiveness of multiplayer online games for children. In *Proceedings of the 6th international conference on Interaction design and children* (pp. 29-36). ACM.

Gerling, K. M., Klauser, M., & Niesenhaus, J. (2011, September). Measuring the impact of game controllers on player experience in FPS games. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 83-86). ACM.

Gerling, K. M., Schulte, F. P., & Masuch, M. (2011, November). Designing and evaluating digital games for frail elderly persons. In *Proceedings of the 8th international conference on advances in computer entertainment technology* (p. 62). ACM.

Gerling, K., & Masuch, M. (2011, June). When gaming is not suitable for everyone: Playtesting Wii games with frail elderly. In *1st Workshop on Game Accessibility*.

Goldsmith, J. T. T., & Ray, M. E. (1948). *U.S. Patent No. 2,455,992*. Washington, DC: U.S. Patent and Trademark Office.

Guo, Y. R., & Goh, D. H. L. (2016, April). From storyboard to software: user evaluation of an information literacy game. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (pp. 199-201). ACM.

Guo, Y. R., & Goh, D. H. L. (2016, December). Heuristic evaluation of an information literacy game. In International Conference on Asian Digital Libraries (pp. 188-199). Springer, Cham.

Gutwin, C., Rooke, C., Cockburn, A., Mandryk, R. L., & Lafreniere, B. (2016, May). Peak-end effects on player experience in casual games. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5608-5619). ACM.

Hara, M., & Ovasaka, S. (2014, October). Heuristics for motion-based control in games. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (pp. 697-706). ACM.

Hartson, R., & Pyla, P. S. (2012). *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, *25*(2), 91-97.

Hochleitner, C., Hochleitner, W., Graf, C., & Tscheligi, M. (2015). A heuristic framework for evaluating user experience in games. In *Game User Experience Evaluation* (pp. 187-206). Springer, Cham.

Hynninen, T. (2012). First-person shooter controls on touchscreen devices: A heuristic evaluation of three games on the iPod touch.

Ijsselsteijn, W. A., De Kort, Y. A. W., & Poels, K. (2008). The game experience questionnaire. *Manuscript in preparation*.

Ijsselsteijn, W., De Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, F. (2007, June). Characterizing and measuring user experiences in digital games. In *International conference on advances in computer entertainment technology* (Vol. 2, p. 27).

Inal, Y., & Cagiltay, K. (2007). Flow experiences of children in an interactive social game environment. *British Journal of Educational Technology*, *38*(3), 455-464.

Isbister, K., & Schaffer, N. (2008). *Game usability: Advancing the player experience*. CRC Press.

ISO, S. (1998). 9241-11 (1998). *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)–Part II Guidance on Usability*.

Järvinen, A. (2005, June). Theory as Game: Designing the Game Game. In *DiGRA Conference*.

Järvinen, A., Heliö, S., & Mäyrä, F. (2002). *Communication and community in digital entertainment services. Prestudy Research Report*.

Jegers, K. (2008, June). Investigating the applicability of usability and playability heuristics for evaluation of pervasive games. In *Internet and Web Applications and Services, 2008. ICIW'08. Third International Conference on* (pp. 656-661). IEEE.

Jegers, K. (2009). *Pervasive GameFlow: Identifying and exploring the mechanisms of player enjoyment in pervasive games* (Doctoral dissertation, Umeå Universitet, Inst för Informatik).

Jerzak, N., & Rebelo, F. (2014, June). Serious games and heuristic evaluation–the cross-comparison of existing heuristic evaluation methods for games. In *International Conference of Design, User Experience, and Usability* (pp. 453-464). Springer, Cham.

Johan, H. (1950). Homo ludens: A study of the play element in culture.

Johnson, D., & Wiles, J. (2003). Effective affective user interface design in games. *Ergonomics*, *46*(13-14), 1332-1345.

Karat, C. M. (1994, June). A comparison of user interface evaluation methods. In *Usability inspection methods* (pp. 203-233). John Wiley & Sons, Inc.

Keller, J. M. (1999). Using the ARCS motivational process in computer-based instruction and distance education. *New directions for teaching and learning*, *1999*(78), 37-47.

Kent, S. L. (2001). The Ultimate History of Video Games: From Pong to Pokémon and Beyond-The Story That Touched Our Lives and Changed the World. Roseville.

Kent, S. L. (2001). *The first quarter: A 25-year history of video games*. BWD Pr.

Khanana, K., & Law, E. L. C. (2013, April). Designing children's digital games on nutrition with playability heuristics. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1071-1076). ACM.

Kiili, K., & Lainema, T. (2008). Foundation for measuring engagement in educational games. *Journal of Interactive Learning Research*, *19*(3), 469.

Koeffel, 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 *Evaluating user experience in games* (pp. 233-256). Springer, London.

Köffel, C., & Haller, M. (2008, April). Heuristics for the evaluation of tabletop games. In *Evaluating User Experiences in Games, Workshop at the 2008 Conference on Human Factors in Computing Systems (Florence, Italy, 5–10 April 2008). CHI* (Vol. 8).

Kokil, U. and Sanchez, J.L.G. (2015). "Exploring facets of playability: the differences between PC and Tablet gaming". 8th International Conference on Advances in Computer-Human Interactions, pp.108-111, Lisbon, Portugal.

Kopomaa, T. (2000). *The city in your pocket: Birth of the mobile information society*. Gaudeamus.

Korhonen, H. (2010, September). Comparison of playtesting and expert review methods in mobile game evaluation. In *Proceedings of the 3rd International Conference on Fun and Games* (pp. 18-27). ACM.

Korhonen, H. (2011, November). The explanatory power of playability heuristics. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology* (p. 40). ACM.

Korhonen, H. (2016). Evaluating playability of mobile games with the expert review method.

Korhonen, H., & Koivisto, E. M. (2006, September). Playability heuristics for mobile games. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 9-16). ACM.

Korhonen, H., & Koivisto, E. M. (2007, September). 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). ACM.

Korhonen, H., Montola, M., & Arrasvuori, J. (2009, October). Understanding playful user experience through digital games. In *International Conference on Designing Pleasurable Products and Interfaces* (Vol. 2009).

Korhonen, H., Saarenpää, H., & Paavilainen, J. (2008). Pervasive Mobile Games–A New Mindset for Players and Developers. In *Fun and Games* (pp. 21-32). Springer, Berlin, Heidelberg.

Kotler, P. (1993). The major tasks of marketing management. *Marketing Management*, *2*(3), 52.

Kultima, A. (2009, September). Casual game design values. In *Proceedings of the 13th international MindTrek conference: Everyday life in the ubiquitous era* (pp. 58-65). ACM.

Lapaš, T., & Orehovački, T. (2015, August). Evaluation of user experience in interaction with computer games. In *International Conference of Design, User Experience, and Usability* (pp. 271-282). Springer, Cham.

Lauer, H. C., & Needham, R. M. (1979). On the duality of operating system structures. ACM SIGOPS Operating Systems Review, 13(2), 3-19.

Laux, L. F. (2001). Aging, communication, and interface design. *Communication, technology and aging: Opportunities and challenges for the future*, 153-168.

Lazzaro, N. (2008). Why We Play: Affect and the Fun of Games: Designing Emotions for Games, Entertainment Interface and Interactive Products. *The Human-Computer Interaction Handbook*, 679-700.

Lee, K. M., & Peng, W. (2006). What do we know about social and psychological effects of computer games? A comprehensive review of the current literature. *Playing video games: Motives, responses, and consequences*, 327-345.

Liao, Y. H., & Shen, C. Y. (2012, March). Heuristic evaluation of digital game based learning: a case study. In *Digital Game and Intelligent Toy Enhanced Learning (DIGITEL), 2012 IEEE Fourth International Conference on* (pp. 192-196). IEEE.

Livingston, I. J., Mandryk, R. L., & Stanley, K. G. (2010, May). Critic-proofing: how using critic reviews and game genres can refine heuristic evaluations. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology* (pp. 48-55). ACM.

M. B. Barbosa, A. B. Rêgo, I. De Medeiros. (2015) HEEG: Heuristic Evaluation for Educational Games. In: Proceedings of SBGames 2015, pp. 224-227

Majgaard, G., & Jessen, C. (2009). Playtesting the digital playground. In *The IADIS International Conference Games and Entertainment Technologies*. International Association for Development, IADIS.

Malone, T. W. (1982, March). 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). ACM.

Marciano, J. N., Miranda, L. C. D., & Miranda, E. E. C. D. (2014). Evaluating multiple aspects of educational computer games: literature review and case study. *International Journal of Computer Games Technology, 2014*, 14.

Mayra, F. (2008). *An introduction to game studies*. Sage Publications Ltd.

Mirza-Babaei, P., Long, S., Foley, E., & McAllister, G. (2011, September). Understanding the Contribution of Biometrics to Games User Research. In *DiGRA Conference*.

Mirza-Babaei, P., Moosajee, N., & Drenikow, B. (2016, October). Playtesting for indie studios. In *Proceedings of the 20th International Academic MindTrek Conference* (pp. 366-374). ACM.

Mohamed, H., & Jaafar, A. (2013). Prioritization of Playability Heuristic Evaluation for Educational Computer Games (Pheg) Technique Using Analytic Hierarchy Process. In *Proceedings of the International Symposium on the Analytic Hierarchy Process 2013*.

Muller, M. J., & McClard, A. (1995, May). Validating an extension to participatory heuristic evaluation: quality of work and quality of work life. In *Conference companion on Human factors in computing systems* (pp. 115-116). ACM.

Nacenta, M. A., Pinelle, D., Stuckel, D., & Gutwin, C. (2007, May). The effects of interaction technique on coordination in tabletop groupware. In Proceedings of Graphics interface 2007(pp. 191-198). ACM.

Nacke, L. (2009). *Affective ludology: Scientific measurement of user experience in interactive entertainment* (Doctoral dissertation, Blekinge Institute of Technology).

Nacke, L. E., Drachen, A., & Göbel, S. (2010). Methods for evaluating gameplay experience in a serious gaming context. *International Journal of Computer Science in Sport, 9*(2), 1-12.

Nacke, L. E., Stellmach, S., Sasse, D., & Lindley, C. A. (2010). Gameplay experience in a gaze interaction game. *arXiv preprint arXiv:1004.0259*.

Nacke, L., Schild, J., & Niesenhaus, J. (2010). Gameplay experience testing with playability and usability surveys–An experimental pilot study. In *Proceedings of the Fun and Games 2010 Workshop, NHTV Expertise Series* (Vol. 10).

Nattkemper, D., & Prinz, W. (1990). Local and global control of saccade amplitude and fixation duration in continuous visual search.

Nielsen, J. (1994). *Usability engineering*. Elsevier.

Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.

Nielsen, J. (1995). Severity ratings for usability problems. *Papers and Essays*, *54*, 1-2.

Nielsen, J., & Molich, R. (1990, March). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 249-256). ACM.

Nielsen, J., & Phillips, V. L. (1993, May). Estimating the relative usability of two interfaces: heuristic, formal, and empirical methods compared. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 214-221). ACM.

Nordin, A. I., Denisova, A., & Cairns, P. (2014, October). Too many questionnaires: measuring player experience whilst playing digital games. In *Seventh York Doctoral Symposium on Computer Science & Electronics* (Vol. 69).

Norman, D. A. (1988). The Design of Everyday Things; First MIT Press: New York, USA.

Norman, D. A. (1991). Cognitive artifacts. *Designing interaction: Psychology at the human-computer interface*, *1*, 17-38.

Oksanen, K. (2013). Subjective experience and sociability in a collaborative serious game. *Simulation & Gaming*, *44*(6), 767-793.

Omar, H. M., & Jaafar, A. (2010). Heuristics evaluation in computer games. In *International Conference on Information Retrieval and Knowledge Management: Exploring the Invisible World, CAMP'10*.

Paavilainen, J. (2010, May). 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* (pp. 56-65). ACM.

Pagulayan, R. J., Steury, K. R., Fulton, B., & Romero, R. L. (2003). Designing for fun: User-testing case studies. In *Funology* (pp. 137-150). Springer, Dordrecht.

Pagulayan, R., & Steury, K. (2004). Beyond usability in games. *Interactions*, *11*(5), 70-71.

Papaloukas, S., Patriarcheas, K., & Xenos, M. (2009, September). Usability assessment heuristics in new genre videogames. In *Informatics, 2009. PCI'09. 13th Panhellenic Conference on* (pp. 202-206). IEEE.

Pinelle, D., & Gutwin, C. (2008). Evaluating teamwork support in tabletop groupware applications using collaboration usability analysis. *Personal and Ubiquitous Computing*, *12*(3), 237-254.

Pinelle, D., Wong, N., & Stach, T. (2008, April). Heuristic evaluation for games: usability principles for video game design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1453-1462). ACM.

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009, May). Usability heuristics for networked multiplayer games. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 169-178). ACM.

Poels, K., De Kort, Y., & Ijsselsteijn, W. (2012). Identification and categorization of digital game experiences: a qualitative study integrating theoretical insights and player perspectives. *Westminster Papers in Communication and Culture*, *9*(1).

Ponnada, A., & Kannan, A. (2012, August). Evaluation of mobile games using playability heuristics. In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics* (pp. 244-247). ACM.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). Human-Computer Interaction. Wokingham, England, Reading, Mass., Menlo Park. *Cal.: Addison-Wesley*.

Quinn, M. M., Symborski, C., Barton, M., Korris, J., Falstad, T., & Granato, S. (2013, September). Methodology for playtesting serious games A case study using a mixed method approach. In *Games Innovation Conference (IGIC), 2013 IEEE International* (pp. 222-227). IEEE.

Raffaele, R., Alencar, R., Júnior, I., Colley, B., Pontes, G., Carvalho, B., & Soares, M. M. (2015, August). Doctor Who: legacy, an analysis of usability and playability of a multi-platform game. In *International Conference of Design, User Experience, and Usability* (pp. 283-291). Springer, Cham.

Reichheld, F. F. (2003). The one number you need to grow. Harvard business review, 81(12), 46-55.

Reichheld, F. F., & Markey, R. (2011). *The ultimate question 2.0: How net promoter companies thrive in a customer-driven world*. Harvard Business Press.

Reimer, J. (2005). Cross-platform game development and the next generation of consoles. *Ars Technica*, *7*.

Röcker, C., & Haar, M. (2006, May). Exploring the usability of videogame heuristics for pervasive game development in smart home environments. In *Proceedings of the third international workshop on pervasive gaming applications–pergames* (pp. 199-206).

Rodio, F., & Bastien, J. M. (2013, November). Heuristics for Video Games Evaluation: How Players Rate Their Relevance for Different Game Genres According to Their Experience. In *Proceedings of the 25th Conference on l'Interaction Homme-Machine* (p. 89). ACM.

Rogers, Y., Sharp, H., Preece, J., & Tepper, M. (2007). Interaction design: beyond human-computer interaction. *netWorker: The Craft of Network Computing*, *11*(4), 34.

Rouse, R. (2001). Games on the verge of a nervous breakdown: emotional content in computer games. *ACM SIGGRAPH Computer Graphics*, *35*(1), 6-10.

Sáenz-de-Urturi, Z., García Zapirain, B., & Méndez Zorrilla, A. (2015). Elderly user experience to improve a Kinect-based game playability. *Behaviour & Information Technology*, *34*(11), 1040-1051.

Sánchez, J. G., Zea, N. P., & Gutiérrez, F. L. (2009, August). Playability: How to identify the player experience in a video game. In *IFIP Conference on Human-Computer Interaction* (pp. 356-359). Springer, Berlin, Heidelberg.

Sánchez, J. L. G., Vela, F. L. G., Simarro, F. M., & Padilla-Zea, N. (2012). Playability: analyzing user experience in video games. *Behaviour & Information Technology*, *31*(10), 1033-1054.

Sánchez, J. L. G., Zea, N. P., & Gutiérrez, F. L. (2009, July). From usability to playability: Introduction to player-centered video game development process. In *International Conference on Human Centered Design* (pp. 65-74). Springer, Berlin, Heidelberg.

Schaffer, N. (2007). Heuristics for usability in games. *White Paper*, 1-30.

Schaffer, N. (2008). Heuristic evaluation of games. *K. Isbister and N. Shaffer, Game Usability. Morgan Kaufman, Amsterdam et al*, 79-89.

Schell, J. (2008). *The art of game design: A deck of lenses*. Schell Games.

Seager, T. P., Lambert, J. H., & Gardner, K. H. (2007). Fostering innovation in contaminated sediments management through multicriteria technology assessment and public participation. *Risk analysis*, *27*(4), 1043-1052.

Seif El-Nasr, M., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H., & Mah, S. (2010, April). Understanding and evaluating cooperative games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 253-262). ACM.

Smeddinck, J. D., Mandryk, R. L., Birk, M. V., Gerling, K. M., Barsilowski, D., & Malaka, R. (2016, May). How to present game difficulty choices?: Exploring the impact on player experience. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5595-5607). ACM.

Song, S., & Lee, J. (2007). RETRACTED: Key factors of heuristic evaluation for game design: Towards massively multi-player online role-playing game.

Soomro, S., Ahmad, W. F. W., & Sulaiman, S. (2012, June). A preliminary study on heuristics for mobile games. In *Computer & Information Science (ICCIS), 2012 International Conference on* (Vol. 2, pp. 1030-1035). IEEE.

Soomro, S., Ahmad, W. F. W., & Sulaiman, S. (2013, November). Evaluation of mobile games using playability heuristics. In *International Visual Informatics Conference* (pp. 264-274). Springer, Cham.

Suhonen, K., & Väätäjä, H. (2010, October). Assessing the applicability of modular playability heuristics for evaluating health-enhancing games. In *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 147-150). ACM.

Sutcliffe, A., & Hart, J. (2017). Analyzing the role of interactivity in user experience. *International Journal of Human–Computer Interaction*, *33*(3), 229-240.

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, *3*(3), 3-3.

Takatalo, J. (2002). Presence and flow in virtual environments: An explorative study. *Master's thesis, University of Helsinki*.

Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, involvement, and flow in digital games. In *Evaluating user experience in games* (pp. 23-46). Springer, London.

Tan, C. T., & Johnston, A. (2011, January). Towards a Non-Disruptive, Practical and Objective Automated Playtesting Process. In *Artificial Intelligence in the Game Design Process*.

Tan, C. T., Rosser, D., Bakkes, S., & Pisan, Y. (2012, July). A feasibility study in using facial expressions analysis to evaluate player experiences. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System* (p. 5). ACM.

Tan, J. L., Goh, D. H. L., Ang, R. P., & Huan, V. S. (2010, October). Usability and playability heuristics for evaluation of an instructional game. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 363-373). Association for the Advancement of Computing in Education (AACE).

Thomas, C., & Milan, S. (1987). Which Input Device Should Be Used With Interactive Video? *In Human–Computer Interaction–Interact'87* (Pp. 587-592).

Tsekleves, E., Cosmas, J., Aggoun, A., & Loo, J. (2009). Converged digital TV services: The role of middleware and future directions of interactive television. *International Journal of Digital Multimedia Broadcasting*, *2009*.

Tsekleves, E., Whitham, R., Kondo, K., & Hill, A. (2009, June). Bringing the television experience to other media in the home: an ethnographic study. In *Proceedings of the 7th European Conference on Interactive TV and Video* (pp. 201-210). ACM.

Tullis, T., Albert, W., Dumas, J. S., & Loring, B. A. (2008). Measuring the User Experience: Collecting Analyzing, and Presenting Usability.

Ülger, G. (2013). Playability heuristics for mobile games using touchscreen displays (Doctoral dissertation, M. Sc. dissertation, Middle East Technical University).

Warren, S., Jones, G., & Lin, L. (2011). Usability and play testing. In *Serious educational game assessment* (pp. 131-146). SensePublishers.

Winn, B., & Heeter, C. (2006). Resolving conflicts in educational game design through playtesting. *Innovate: Journal of Online Education*, *3*(2).

Wixon, D., & Pagulayan, R. (2008). THAT'S ENTERTAINMNT Halo 3: the theory and practice of a research-design partnership. *Interactions*, *15*(1), 52-55.

Wodike, O. A., Sim, G., & Horton, M. (2014, September). Empowering teenagers to perform a heuristic evaluation of a game. In *Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014-Sand, Sea and Sky-Holiday HCI* (pp. 353-358). BCS.

Wood, E., Willoughby, T., Rushing, A., Bechtel, L., & Gilbert, J. (2005). Use of computer input devices by older adults. Journal of Applied Gerontology, 24(5), 419-438.

Wood, W., Harris, S., Snider, M., & Patchel, S. A. (2005). Activity situations on an Alzheimer's disease special care unit and resident environmental interaction, time use, and affect. *American Journal of Alzheimer's Disease & Other Dementias®*, *20*(2), 105-118.

Zaibon, S. B., & Shiratuddin, N. (2010, April). Heuristics evaluation strategy for mobile game-based learning. In *Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE), 2010 6th IEEE International Conference on* (pp. 127-131). IEEE.

Zaman, L., Natapov, D., & Teather, R. J. (2010, May). Touchscreens vs. traditional controllers in handheld gaming. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology* (pp. 183-190). ACM.

Zhao, C., Zhang, T., & Zhang, K. (2001, October). User interface design for the small screen display. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 45, No. 22, pp. 1548-1550). Sage CA: Los Angeles, CA: SAGE Publications.

Zhu, M., Zhao, F., Fang, X., & Moser, C. (2017). Developing Playability Heuristics Based on Nouns and Adjectives from Online Game Reviews. *International Journal of Human–Computer Interaction*, *33*(3), 241-253.

Ziefle, M. (2010). Information presentation in small screen devices: The trade-off between visual density and menu foresight. *Applied ergonomics*, *41*(6), 719-730.

Ziefle, M., & Bay, S. (2005). How older adults meet complexity: aging effects on the usability of different mobile phones. *Behaviour & Information Technology*, *24*(5), 375-389.

Zimmermann, E., & Salen, K. (2004). Rules of play: Game design fundamentals.

Zook, A., Fruchter, E., & Riedl, M. O. (2014). Automatic playtesting for game parameter tuning via active learning. In *FDG*.

***Electronic Sources***

Gameindustry web article, 2017. “GamesIndustry.biz presents... The Year In Numbers 2017”, James Batchelor. https://www.gamesindustry.biz/articles/2017-12-20-gamesindustry-biz-presents-the-year-in-numbers-2017

# APPENDICES

## Appendix-A. Results of device-dependent playability problems

**Table A1.** Results of device-dependent playability problems

| Heuristic Item | *Exemplary Statement* | Main Observations |
|---|---|---|
| **GU1a: "Audio-visual representation supports the game"** | - *"I loved the graphics and animations but the sound effects were repetitive and reedy. Especially the sound of a zombie eating a plant. This annoyed me after certain amount of time" (PM8)* | - Sound effects of the game annoyed mobile players<br>- Mobile players accidentally blocked the speakers of the device<br>- Mobile platform speakers were not sufficient |
| **GU4: "Indicators are visible"** | | - Players' perception for indicators can potentially be related with the screen size of the mobile platform<br>- Mobile players are inspected to perceive indicators easily compared to PC players<br>- Mobile platforms require less cognitive, spatial or attentional demand |
| **GU7: "Game controllers are consistent and follow standard conventions" & GU8: "Game controls are convenient and flexible"** | - *"I would prefer a drag mechanism to play this game" (PM9)* | - Both player groups found the controls sufficient and convenient in general<br>- Even though these heuristics items included the notion of controls being flexible for different preferences of players, the game did not provide such an option. Nevertheless players gave higher scores to these items.<br>- Heuristics may not clearly be understood by the players<br>- The wordings of these heuristics did not relate to the fundamental control differences between platforms in terms of usability |
| **GU11: "The player does not have to memorize things unnecessarily"** | - *"…for instance now I received the Cherry Bomb plant, yet the info they gave me is only limited to its basic application. I would really prefer to see some kind of detailed information" (PP6).* | - Two PC players indicated that they were not able to see the info about the given plants whenever they wanted to and got frustrated about using them<br>- Mobile players experienced the same game design in this manner but they all pointed out that the game was giving them the chance to learn by trial and error<br>- This heuristics is found to be in close relationship with both the heuristic item *GU4* and the ease of use of the mobile platform.<br>- Mobile platforms require less cognitive, spatial or attentional demand |
| **GP4: "Player is in control"** | - *"Although I knew that I miss-tapped from time to time, I never felt that I lost the control of the game" (PM1).* | - Mobile platform offered the capability to manipulate the object on the screen directly by either touching or dragging, reducing the cognitive load while PC platform only allowed the players to interact with a mouse via clicking to the objects<br>- Players on mobile platform had several problems interacting with the game such as tapping on the wrong plants, they indicated that they felt in control throughout the game. |

| | | |
|---|---|---|
| | | - This contrast between errors and players feeling of direct control on mobile platform may become more significant compared to the context of usability and cause a higher feeling of control hence more error tolerance. |
| **MO2: "The game accommodates the surroundings"** | *"I was not happy with the music of the game, but I was able to lower the volume fairly easy" (PM4).* | - Players were able to adjust the volume easily during the play.<br>- Several of the players (n=3) were observed to lower the master volume of the device by utilizing the device volume button rather than the in-game options menu.<br>- Players considered the heuristic item was referring to device capabilities rather than the game. |
| **MO4: "The graphical design is accommodated to current brightness"** | | - The item has a strong relationship with the *MO2* & *MO3* heuristic items in terms of context.<br>- Players indicated that the screen brightness is usually about the device technology.<br>- The results suggested that the players did not relate the item with *GU1a* but with the device capabilities |
| **MO6: "Mobile devices have their own conventions for input"** | | - Players gave higher scores for the item on average.<br>- Two players (PM1 & PM10) had difficulty understanding the heuristic and hesitated to answer either positively or negatively.<br>- This item is given as a support for *GU7* heuristic item. This association is validated through observations. The players which gave higher scores indicated that the conventions for the mobile platform were obvious and natural. |

**Table A2.** Results of design-dependent playability problems

| | | |
|---|---|---|
| **GU2: "Screen layout is efficient and visually pleasing"** | - *"I accidentally pressed the menu button of the device during play and it paused the game and dropped me out from the game. I had this issue before but I just realized it was because of the devices menu button" (PM3)* | - Two PC players preferred higher resolution<br>- Mobile players accidentally tapped operating system menu button and dropped out of the game<br>- All of the players (n=20) liked the aesthetics of the game |
| **GU3: "Device UI and game UI are used for their own purposes"** | - *"I was trying to select a plant but instead I accidentally pushed to the Android system menu button on the bottom of the screen by accident. It kicked me out of the game for a second" (PM9).* | - PC players clearly understood the difference between Game UI and the UI of the operating system<br>- The accidental drop-outs from the game on mobile platform because of the accidental tap of the menu button of the device |
| **GU5: "The player understands the terminology"** | - *"The story did not immerse me or gained my attention much since I did not feel necessarily invested in it" (PM10).*<br><br>- *"I paused the game during play to see if I was able to read the information about the plants. Unfortunately, the game did not gave me that opportunity" (PM1).*<br><br>- *"I would really prefer if they have told the functions of the plants with some kind of animation instead of bulk text, giving minimal detail" (PM5).* | - Mobile players skipped the information screens including plant informations and story elements<br>- None of the PC players skipped the information screens during the game.<br>- Mobile players did not understand some of the functions of the plants in the game<br>- Mobile players wanted to re-check plant information with no success |
| **GU6: "Navigation is consistent, logical and minimalist"** | - *"The buttons were embedded in the background and were not visible. I had to ask for help to begin the game" (PM3).* | - Mobile players found the navigation problematic<br>- Mobile players had severe stagnation in the main menu screen at the beginning of the game.<br>- Mobile players could not find the button to start the game immediately.<br>- Two mobile players had to ask how to start the game to the supervisor.<br>- PC version of the game was different in terms of providing highlighted button effects, making the buttons potentially more distinct than the background image. |
| **GP1: "The game provides clear goals or supports player-created goals"** | - *"I did not know that the Tower Defense games were also made for mobile phones" (PM6).* | - The tutorials at the beginning of the game provided sufficient information related to the mechanics yet did not provide information of long-term goals of the game.<br>- PC platform players understood the necessary goals and ways to achieve them more clearly than mobile players<br>- The familiarity with the *tower defense* game genre may have an impact regarding this heuristic<br>- Two mobile players (PM6, PM9) indicated that they have never played the game genre on a mobile phone during the game |
| **GP2: "The player sees the progress in the game and can compare results"** | - *"The high volume level kind of bothered me but I don't want to pause the game and lower it now because I am following the progress bar at the bottom of the screen and know that the level is going to end soon" (PP8).*<br><br>- *"I think I won't be able to tell any of my friends on which level I was at if need be because I really don't understand what these level numbers indicate" (PM5).* | - On mobile platform, the progress bar is presented on top of the screen while and on PC, it is placed at the bottom of the screen<br>- Mobile players have too many focal points on top of the screen.<br>- Progress bar on mobile platform is next to several other game elements such as the menu button while PC progress bar is isolated from other in-game objects, hence more visible. |

*Table B1*

| | | |
|---|---|---|
| **GP5: "Challenge, strategy and pace are in balance"** | - *"I purposefully tried to lose the game because I was bored of not being challenged. But still I managed to win it without and hassle" (PM3).*<br><br>- *"I think the mobile platforms are best for fast-paced games and this game is truly offering the opposite" (PM4).*<br><br>- *"I am used to playing faster-paced games on my mobile phone because I usually want to enjoy the game in a short time" (PM6).* | - Players from both platforms complained about the pacing of the game and its lack of challenge<br>- Mobile players especially commented on the lack of challenge and slow pace.<br>- This heuristic was observed to be a device dependent item<br>- PC version of the game is perceived to be less cumbersome compared to the mobile platform |
| **GP6: "The first-time experience is encouraging"** | | - This heuristic is related to the onboarding process<br>- A mobile-first approach is suggested to be more beneficial regarding this heuristic<br>- First-time experience for mobile players was problematic because of the start menu problems.<br>- This item is associated with *GP1* |
| **GP7: "The game story, if any, supports the gameplay and is meaningful"** | - *"This would be the same game even if you put elephants for instance, instead of zombies and turrets instead of plants" (PM10).* | - All PC players indicated that the story of the game felt fun and enjoyable<br>- A Number of mobile players mentioned that they did not understand the story of the game and not found the ambient crucial for the concept of the game (n=2)<br>- Onboarding process differences between platforms is once more observed. The mobile players had difficulty understanding the overall story of the game and purpose of the game characters in terms of context. |
| **GP8: "There are no repetitive or boring tasks"** | - *"I've seen better versions of tower defense genre. This game only offers simple tasks and not enough surprises" (PM7).* | - Mobile players mentioned that they were familiar to the tower defense game genre for mobile platform<br>- The game genre is suitable for the mobile platform yet it lacked some of the important game elements such as sufficient variations of game-related tasks |
| **GP11: "The game does not stagnate"** | - *"Normally I would quit the game if this was not a test environment and I was not curious about the next levels of the game. The pacing is really boring me right now" (PM9).* | - PC players felt less stagnation and more progress<br>- The players from both platforms associated this heuristic item to the pacing of the game.<br>- Heuristic item may identify with the heuristic item *GP5*<br>- Mobile players complained about the slow pace of the game during play sessions. |
| **GP12: "The game is consistent"** | - *"The game informed me by telling that the cherry bomb plant is just like another plant. So I planted it among my defense line yet it exploded immediately! The game did not inform me about this!" (PM5).* | - Players from both platforms understood the heuristic item is to evaluate the relation between plant information and functions of the plants.<br>- Both groups' made similar mistakes during the game by misinterpreting a function of a plant.<br>- Mobile players complained mostly on how the game was not sufficiently explaining the functions of the given plants. |
| **GP13: "The game uses orthogonal unit differentiation"** | - *"I couldn't figure out the function of the 'mine' plant. I tried to use it many times but it didn't do anything. At the end I decided to use it as an emergency zombie blocking plant because it is cheap to purchase" (PM7).* | - The players on mobile platform had difficulty in understanding specific functions and strategic implications of some of the plants.<br>- This heuristic item could be considered to be in association with the heuristic item *GU5* |
| **MO3: "Interruptions are handled reasonably"** | - *"This heuristic item is related to the mobile device that I am using now and not the game itself" (PM6).* | - Players were observed to accidentally press the device menu button. Afterwards, they get back to the game from where they were interrupted immediately by pressing the same button again. They mentioned that they |

| | | |
|---|---|---|
| | | - This heuristic item should be considered in relation with the *GU3* heuristic item were able to return to the game in a fast manner |
| **MO7: "The tutorial should respond to immediate demand"** | - "*I did not recognize such a function in the game" (PM4).* | - The tutorial was not immediate, yet players understood the heuristic was related to having tutorial levels at the beginning.<br>- The item is referred to *GU12* heuristic yet for instance PM6 gave the lowest score to this heuristic item while giving the highest score to *GU12*<br>- The proposed link between these heuristics might not be there when applied in practice |

**Table A3.** Results of game-specific playability problems

| | | |
|---|---|---|
| **GU1b: "A view to the game-world supports smooth interaction and the camera behaves correctly"** | | - **In-game camera is static**<br>- **The game does not have the necessary property, therefore not applicable.** |
| **GU9: "The game gives feedback on the player's actions & GU10: "The player cannot make irreversible errors"** | | - it was identified that these two heuristics out of the context and not applicable for analyzing differences between platforms |
| **GU12: "The game contains help"** | - *"There was not help function as much as I remember" (PM3).* | - The game does not directly provide a help function neither for PC nor mobile platform.<br>- The game provides well designed tutorial levels.<br>- Nearly all of the players mentioned that they did not need any help function during the game, sometimes referring to lack of challenge of the game<br>- Players were relating this heuristic item to the lack of help functionality during the game yet the heuristic item also explained to include tutorials of games |
| **GP3: "The players are rewarded and the rewards are meaningful"** | | - No notable difference was observed.<br>- Other than the progression related rewards, there was no score indicator. Hence no difference between platforms in the context of this heuristic was expected. |
| **GP9: "The players can express themselves" & GP10: "The game supports different playing styles"** | | - Although the game does not allow players to customize their avatars or personal preferences, PC players mentioned the personalization of different strategies during the play sessions might count as expressions.<br>- These heuristic items does not apply for neither of the versions of the game and is not applicable to evaluate platform differences in terms of player experience<br>- *GP10* did not receive any feedback regarding the difference between platforms since the game did not offer any capability to choose a role and/or style of play in the world. There were also no alternative interaction capabilities implemented in the game. |
| **GP14: "The player does not lose any hard-won possessions"** | | - The game for both platforms have the same mechanics and dynamics and fulfill this heuristic sufficiently |
| **MO1: "The play sessions can be started quickly"** | | - One player (PM9) was observed to tap on the screen while the game was first opening on the device.<br>- Most of the mobile players were referring this heuristic to the device capabilities. |

# CURRICULUM VITAE

Çakır AKER
PhD Candidate, Media and Communication Studies
Galatasaray University
cakiraker@gmail.com

## Education

2013 –2018 Ph.D. in Media and Communication Studies
Graduate School of Social Sciences, Galatasaray University, Turkey.

2008 – 2010 Postgraduate Master Degree in Information Technologies, Communication & Audiovisual Media, Interdisciplinary Master of Cognitive Systems and Interactive Media, University of Pompeu Fabra, Spain.
Thesis: We Hunters: Analyzing the Educational Value of A Virtual Heritage Experience Based On A Multi-Touch Screen (2010)

2006 – 2008 Postgraduate Master Degree in Graduate School of Social Sciences, Strategic Marketing and Brand Management, Bahcesehir University, Turkey.

2001 – 2005 University Degree in Advertising, Faculty of Communication, Bahcesehir University, Turkey.

## Employment

2012 – Present Research Assistant at Bahcesehir University, Faculty of Communication, Department of Digital Game Design

| | |
|---|---|
| 2014 – Present | Coordinator at Bahcesehir University Game Lab, Department of Digital Game Design. |
| 2010 – 2011 | Interface Designer at University of Pompeu Fabra and Barcelona Media. |
| 2009 – 2010 | Advertising Consultant at Doğan Automotive, Freelance Contractor |
| 2007 – 2008 | General Coordinator at Best Image Advertising Agency |
| 2005 – 2007 | Copywriter at Alaaddin Advertising Agency |

**Academic Works**

Kourakis Stelios, Aker Çakir, Parés Narcís (2012). We hunters: interactive communication for young cavemen. In *International Journal of Arts and Technology*, (pp. 199-220). Springer.

Aker Ç., Öztürk Ö. (2015) Social Media Interactions and the Use of Third-Party Management Applications on Effectiveness and Perception of Information. In: Marcus A. (eds) Design, User Experience, and Usability: Interactive Experience Design. DUXU 2015. Lecture Notes in Computer Science, vol 9188. Springer, Cham

Aker Ç., Rızvanoğlu K., İnal Y., Yılmaz A.S. (2016) Analyzing Playability in Multi-platform Games: A Case Study of the Fruit Ninja Game. In: Marcus A. (eds) Design, User Experience, and Usability: Novel User Experiences. DUXU 2016. Lecture Notes in Computer Science, vol 9747. Springer, Cham

Aker, Ç., Rızvanoğlu, K. & İnal, Y. (2017). A Multi-Modal Approach for Evaluating Player Experience on Different Gaming Platforms. In Press

Aker, Ç., Rızvanoğlu, K., & Bostan, B. (2018) Methodological Review Of Playability Heuristics. In Press

**TEZ ONAY SAYFASI**

| | |
|---|---|
| Üniversite | : T.C. GALATASARAY ÜNİVERSİTESİ |
| Enstitü | : SOSYAL BİLİMLER ENSTİTÜSÜ |
| Hazırlayanın Adı Soyadı | : Çakır AKER |
| Tez Başlığı | : Analyzing Player Experience: Multi-modal Evaluation of Player Experience in Different Gaming Platforms |
| Savunma Tarihi | : 28 / 06 / 2018 |
| Danışmanı | : Prof. Dr. Kerem RIZVANOĞLU |

JÜRİ ÜYELERİ

| Unvanı, Adı Soyadı | İmza |
|---|---|
| Prof. Dr. Kerem RIZVANOĞLU | |
| Prof. Dr. Serhat GÜNEY | |
| Doç. Dr. Barbaros BOSTAN | |
| Dr. Öğr. Üyesi Güven ÇATAK | |
| Dr. Öğr. Üyesi Ece VİTRİNEL | |

Enstitü Müdürü
Prof. Dr. M. Yaman ÖZTEK