Methodological Review Of Playability Heuristics

Çakır Aker
Faculty of Communication
Department of Digital Game Design
Bahcesehir University
Istanbul, Turkey
cakir.aker@comm.bau.edu.tr

Kerem Rızvanoğlu
Faculty of Communication
Galatasaray University
Istanbul, Turkey
krizvanoglu@gsu.edu.tr

Barbaros Bostan
Faculty of Communication
Department of Digital Game Design
Bahcesehir University
Istanbul, Turkey
barbaros.bostan@comm.bau.edu.tr

Abstract— **This study reviews published scientific literature on the use of heuristics for evaluating player experience to (a) identify the potential contribution of the application of heuristics for assessing player experience, (b) present the status of playability heuristics evaluation procedures for evaluating player experience, and (c) define future research perspectives. By searching online bibliographic databases, 44 relevant articles were selected and included in the study. The aim, methodology, proposed heuristics and conclusions were studied separately for each article. The study indicated that a large variety of approaches on evaluating the experience of video games using heuristics were presented. Because of this, it is not possible to identify a generally accepted approach while studying evaluation of video games. This study intends to present and clarify a much-needed holistic point of view in terms of using heuristics for evaluating player experience because of the current dispersed state of the literature and interlaced heuristic evaluation approaches. The review study indicated that most of the articles presented new heuristics, either by iteratively improving the existing approaches or forming new ones. It is usually suggested that the presented sets of heuristics have been viable in general for assessing some aspects regarding the gaming experience. However, the heuristic approaches have neither proven to be including the experience in its entirety nor empirically tested adequately for validation to provide a possible de facto basis for further research. The implications of the articles were also studied for providing a common ground for future research in the field of heuristics evaluation of video games.**

Keywords—playability; playability heuristics; heuristic evaluation; literature review; game experience

I. INTRODUCTION

Most of the usability evaluation techniques are not suitable for inspecting video games since the design considerations and aims for games are different from productivity applications. While productivity applications focus on solving problems and minimizing challenges, video games use challenges and problems to provide enjoyment. More so, video games do not have specific tasks to achieve a specific goal but depend on player choices and motivations. In order to satisfy user needs and enhance the overall experience for both the productivity and video game applications, there are several user evaluation techniques. One example is task-oriented user tests which the data gathered from a sample of users interacting with the application. Users are observed while following simple assignments representing a typical use of the application. The results are analyzed for indicating issues related to the user experience of the application. Differently, among those methods for evaluating an interactive system, heuristic evaluation offers the benefit of evaluation during the design process and do not require a task oriented inspection.

The overall importance of games, both in terms of industry and academia, has been growing rapidly but the evaluation of player experience has been complicated by the following problems: (1) The literature on player experience and playability is not vast; (2) there is no common definition for playability and there is no consensus on the heuristics for evaluating the playability in video games; and (3) researchers apply different approaches and therefore offer different heuristics for evaluation. In this review, our aim is to present both the goals and the procedures of using heuristics as a means of evaluating playability and player experience in video games. Likewise, we also aim to find out the future potential of various methods utilizing playability heuristics.

The game industry has been developing immensely although the methodology on evaluating player experience still lack a robust approach to evaluate the overall experience. Heuristic evaluation is an inspection technique that allows evaluators to examine an interface using statements of usability principles [1]. Korhonen mentions that heuristic evaluation is more effective for evaluating games compared to other methods because this approach does not require any tasks oriented tests to be conducted [3]. The studies on player experience using heuristics have the benefit of conducting a research with a rather cheap and fast manner. So far, researchers who studied the topic have not been able to present a holistic set of heuristics which could be considered as a common ground for evaluating player experience. Moreover, only a limited number of studies aimed to utilize existing approaches and even fewer attempted to combine different approaches and heuristics from previous works and validate them with empirical tests. Only a few of the researchers have tried to present new playability heuristics [7], [6], [4], [17], [11], [13] for games. Choosing heuristics which

Aker, C., Rizvanoglu, K., & Bostan, B. (2017) Methodological Review Of Playability Heuristics. Proceedings of the Eurasia Graphics 2017, Istanbul, Turkey, November 4-5.

are applicable in a selected gaming platform is a challenging task and some of the researchers have proposed different heuristics for different platforms such as tabletop, computer, mobile, educational and web based social games [11], [13], [16], [20], [12], [22], [28], [30]. In addition, reviewed literature has presented that each heuristic method has pros and cons. This study aims to provide a comprehensive review on using heuristics for evaluating playability to improve the field of research regarding the evaluation of player experience.

II. METHOD

An extensive literature search was conducted with key-words 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 this endeavor, a second examination of the abstracts led to an elimination of irrelevant journal articles and conference proceedings. The remaining 44 articles were examined in detail to find out the main contributions to the relevant research area. The findings of our literature study addressed some of the key differences for presenting heuristics for evaluating player experience. These key features were categorized in terms of what procedure the heuristics were based upon such as the choice of medium or sources for identification. The rest of the paper was structured as follows, first the categorization for procedures in the articles was explained. Four distinctive methods were identified; empirical evaluations, expert evaluations, inspections (literature & online game reviews) and evaluations using mixed methodology. Afterwards, studies in each category were summarized. Explanations of the articles utilizing empirical evaluation methods, expert evaluation methods, inspections and multi-modal methodology are presented followed by conclusions.

III. PROCEDURE TYPES

In terms of the categorization of playability heuristics, several studies presented relevant articles in a chronological manner. Even though this structure for studying literature review has its benefits, such as indicating the iterative progress between testing methods and design methodology, it is not sufficient for examining the procedural differences between approaches. Moreover, our review focuses on structuring the literature in terms of methodology. Because of these reasons, categorizing differences in general between heuristic evaluation methods applied during tests and development of heuristic approaches hold a different perspective for analysis and potential for contribution. As a result, it is possible to conclude that there are four main approaches in the field. Among the eligible 44 articles, we identified 12 articles which would fit in the category of empirical evaluation, 17 articles for expert evaluation, seven articles for inspection and 8 articles for mixed-methodology.

Expert evaluations conducted by utilizing proposed heuristics is the most common approach in the literature, followed by empirical evaluations. Although many of the researchers claimed that the best approach for a valid and robust heuristic set is to combine different evaluation methods while testing them one of the least employed methodology is the mixed-method approach (Figure 1). Articles in each category is explained in a chronological fashion to provide an additional standpoint for indicating the extent of scientific progress.

A. Empirical Evaluation

The first observed approach is empirical evaluation method. This category includes studies conducted through user-testing methods such as surveys, interviews, focus groups and observations with a sample group of minimum 10 participants. According to our review, 12 of the relevant articles chose to evaluate either the heuristics or the games via user-tests rather than directly using the heuristic evaluation sets during the research.

Malone [4] proposed the first heuristics for encouraging the use of games in learning and teaching. He presented a set of heuristics for instructional games and suggested that there were three main heuristics for achieving entertainable interfaces. In the study, three empirical tests were employed to understand what gamers liked with a total number of 81 participants, all from elementary or secondary school students. The tests were conducted upon three games with eight versions each. As a result he proposed the three heuristics categories; challenge, fantasy and curiosity.

Fabricatore et al. [5] followed a different approach by first proposing heuristics and later evaluating it. The model was prepared in order to guide game designers for preparing better games. 53 participants, between the ages of 20 – 30, have attended the tests where they could make comments in every step. After play sessions, semi-structured in-depth interviews were realized. In the light of the findings derived from the play sessions, in-depth interviews and observation notes, the proposed heuristics were iteratively revised. 39 different games from different genres were tested during the study. At the end of each test, participants were interviewed for potential improvements of the heuristics. A hierarchical table was presented as determinants of the playability in games with several sub categories in each.

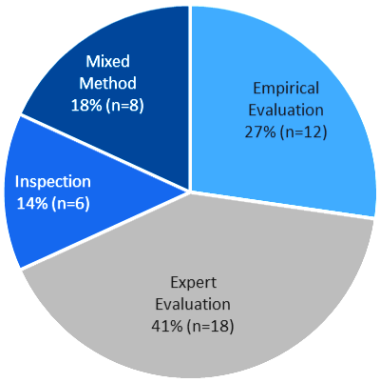

Fig. 1. *Distribution of Different Methods Utilizing Heuristics Evaluation for Playability and Player Experience*

The study indicated three main determinants; entity, scenario and goals. Item of entity consisted of four different sub categories; identity, energy, equipment and behavior. Scenario item consisted of view, spontaneous changes, transitions and interactions with entities. And the last item of hierarchy of goals consisted of complexity, linearity and interface. This study is regarded as one of the first qualitative playability evaluation models presented in the literature, therefore prepared a basis for the rest of the research field.

Inspired by heuristic evaluation approaches Röcker & Haar [7] investigated if the existing heuristics could be used for evaluating pervasive games [10]. Their study was based on Desurvire's HEP heuristics. They conducted the research via a focus group of 10 participants, between the ages of 32 - 38. The evaluated heuristics set was not shown to the participants during the tests for objectivity reasons. A smart home environment with pervasive computing capabilities were given as a scenario for the participants to make comments. After the interviews, participants were invited to join a focus group and asked to note their ideas to cards that were given to them. At the end, they were asked to do a card-sorting exercise to clarify the priorities and categories for each comment. Later, 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 [12] 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 [7] and suggested 54 key factors under four key categories; game interface, gameplay, game narrative, and game mechanics.

Desurvire & Wiberg [19] conducted a research based upon Desurvire's previous approach of using HEP heuristics [7] 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. [22] 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 [51], Keller's ARCS Model for Motivation [52], GameFlow model [9] and Nielsen's heuristics [1]. 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 [30] 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 [11] 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 [37] 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 4 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) [55] and interviews. 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.

Kornchulee Khanana & Effie Lai-Chong Law [38] conducted a study to use the Game Flow [9] 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 [40] conducted a study to evaluate the PLAY heuristics from Desurvire & Wiberg [19]. 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 [41]. 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) [25] 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 used for the application of heuristic evaluation.

Lastly, Guo & Goh [49] 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 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 I. 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

Previous Empirical Evaluation Research on Playability Heuristics	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)
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

B. 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 [1], 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 [6] 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 [1]. As a result, Federoff presented a set of 10 heuristics for evaluating games yet the suggested heuristics lacked any validation.

Baauw, Bekker & Barengregt [8] conducted a study on the proposed Structured Expert Evaluation Method (SEEM) which was inspired from Norman's theory-of-action model [53] and Malone's concepts of fun [4]. 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 [9], 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' [56] 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 [11] 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 [11]. Following their previous work, Korhonen & Koivisto [13] 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 & Haler [16] 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. [18] 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 [13], [7] 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. [20] 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 [17] 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 [59]. 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 [11] 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 [20]. Other researchers criticized the article because the previous work which the heuristics were based on [17] 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. [27] conducted a study to inspect the use of heuristics to evaluate the overall user experience of video 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 [17] and Sweetser & Wyeth's GameFlow approach [9]. 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 [59]. 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. [23] conducted a heuristic evaluation of the web-based computer game 'FarmVille' by combining heuristics from Federoff [6], Desurvire [7] and Pinelle [17]. 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 [28] aimed to study the effect of using modular heuristics on health games. Five previous heuristic sets [7], [6], [61], [11], [13] 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 [11] 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 [25] 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 [33] 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 [11]. 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 [32] 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 [17] 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 [34] followed a rather different approach. Instead of testing the heuristics, he applied a three session Delphi test process [57], [58] 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 [36], conducted a study on refining the previously explained Playability Heuristics Evaluation for Educational Computer Game (PHEG) heuristic set [25]. 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) [62] 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. [42] studied the efficiency of empowering teenagers as expert evaluators for analyzing video games in their paper. Based on Pinelle's [17] 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. [47] conducted a research about heuristic evaluation of educational games, proposing a blend of items from HEP [7], PLAY [19], and GameFlow [9] 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) 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 II. 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		
Ebdoiment of methapors 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 immidiate 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							
Communication and collaboration should be supported by 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 behaviour							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 immidiate positive feedback						X		X		X							

Previous Expert Evaluation Research on Playability Heuristics	Federoff (2002)	Baauw, Bekker & Barengregt (2005)	Sweetser & Wyeth (2005)	Korhonen & Koivisto (2006)	Köffel & Haler (2008)	Korhonen, Paavilainen & Saarenpää (2009)	Pinelle, Wong, Stach & Gutwin (2009)	Koeffel et al. (2010)	Almeida, Mealha, Veloso (2017)	Suhonen & Vaataja (2010)	Omar & Jaafar (2010)	Ponnada & Kannan (2012)	Hynninen (2012)	Carmody (2012)	Mohamed, Jaafar (2013)	Wodlike, Sim, Horton (2014)	Barbosa, Rego, Medeiros (2015)
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 wellcoordinated											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		
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	

C. Inspections

1) Literature Reviews

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

Schaffer [14] 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 [29] 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 [6], [7], [11], and [20]. 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 [11] 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 [45] 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 [6], [7], and [19]. 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) 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. [26] presented a study on using critic reviews of games for refinement of heuristic evaluations. Pinelle's [17] 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 [44] 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. [50] 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. [7], Federoff [6], Malone [4], Pinelle [17], [20] 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 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 III. 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

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)
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

3) Mixed Method

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

Desurvire et al. [7] 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. [17] 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 [21] 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. [54] 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 [1]. 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 [15] studied on defining the enjoyment in pervasive games. Three pervasive tabletop games were tested using the GameFlow model [9] 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 [39] 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 [19] and Game Approachability Principles (GAP) heuristics [48] 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 [48] 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. [46] 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 [27]. 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 [4], Federoff [6], Desurvire [7] Shaffer [14], Pinelle [17], Koeffel [27], and Korhonen & Koivisto [11]. However, the focal point of the suggested new heuristic set was Koeffel's [27] 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 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 IV. 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	

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)
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	
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	

IV. CONCLUSION

By categorization of methodological differences between heuristic evaluation researches in this study, we aimed to present a novel perspective to the domain of playability evaluations. It was also aimed to present 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 also observed that literature review for defining the playability heuristics was the most common way to conduct studies.

The study represented that 15 heuristics were common for all of the methods (Table 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 heuristics 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 V. 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

REFERENCES

[1] J. Nielsen. "Usability inspection methods" In: *Conference companion on Human factors in computing systems*. ACM, 1994, pp. 413-414.

[2] T. Fullerton, C. Swain, S. Hoffman, Game design workshop: Designing, prototyping, & playtesting games. CRC Press, 2004.

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

[4] T. W. Malone, "Heuristics for designing enjoyable user interfaces: Lessons from computer games" In: *Proceedings of the 1982 conference on Human factors in computing systems*. ACM, 1982. pp. 63-68.

[5] C. Fabricatore, M. Nussbaum, R. Rosas, "Playability in action videogames: A qualitative design model" In: *Human-Computer Interaction*, 2002, pp. 311-368.

[6] M. A. Federoff, "Heuristics and usability guidelines for the creation and evaluation of fun in video games," Doctoral Dissertation, Indiana University, 2002.

[7] H. Desurvire, M. Caplan, J. A. Toth, "Using heuristics to evaluate the playability of games" In: *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 2004, pp. 1509-1512.

[8] E. Baauw, M. Bekker, W. Barendregt, "A structured expert evaluation method for the evaluation of children's computer games" In: *Human-Computer Interaction-INTERACT 2005*, 2005, pp. 457-469.

[9] P. Sweetser, P. Wyeth, "GameFlow: a model for evaluating player enjoyment in games," In: *Computers in Entertainment (CIE)*, 2005, v.3 n.3: 3-3.

[10] C. Röcker, M. Haar, "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*. 2006, pp. 199-206.

[11] H. Korhonen, E. Ml. Koivisto, "Playability heuristics for mobile games," In: *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. ACM, 2006, pp. 9-16.

[12] S. Song, J. Lee, Retracted: "Key factors of heuristic evaluation for game design: Towards massively multi-player online role-playing game," 2007, pp. 709-723.

[13] H. Korhonen, E. Ml. Koivisto, "Playability heuristics for mobile multi-player games," In: *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts*. ACM, 2007, pp. 28-35.

[14] N. Schaffer, "Heuristics for usability in games-white paper," 2007.

[15] K. Jegers, "Pervasive GameFlow: Identifying and exploring the mechanisms of player enjoyment in pervasive games," 2009, Doctoral Dissertation, Umeå Universitet, Inst för Informatik.

[16] C. Köffel, M. Haller, "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, 2008.

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

[18] H. Korhonen, J. Paavilainen, H. Saarenpää, "Expert review method in game evaluations: comparison of two playability heuristic sets," In: *Proceedings of the 13th international MindTrek conference: Everyday life in the ubiquitous era*. ACM, 2009, pp. 74-81.

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

[20] D. Pinelle, N. Wong, T. Stach, C. Gutwin, "Usability heuristics for networked multiplayer games," In: *Proceedings of the ACM 2009 international conference on Supporting group work*. ACM, 2009, pp. 169-178.

[21] A. Febretti, F. Garzotto, "Usability, playability, and long-term engagement in computer games," In: *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2009, pp. 4063-4068.

[22] J. L.Tan, D. H. Goh, R. P. Ang, V. S. Huan, "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*. Association for the Advancement of Computing in Education (AACE), 2010, pp. 363-373.

[23] S. Almeida, O. Mealha, A. Veloso, "Heuristic Evaluation of "FarmVille". *VIDEOJOGOS 2010*, 2010.

[24] H. Mohamed, A. Jaafar, "Development and potential analysis of heuristic evaluation for educational computer game (PHEG)," In: *Computer Sciences and Convergence Information Technology (ICCIT), 2010 5th International Conference on*. IEEE, 2010, pp. 222-227.

[25] H. M. Omar, A. Jaafar, "Heuristics evaluation in computer games. In: International Conference on Information Retrieval and Knowledge Management: Exploring the Invisible World, CAMP'10, 2010.

[26] I. J. Livingston, R. L. Mandryk, K. G. Stanley, "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*. ACM, 2010, pp. 48-55.

[27] C. Koeffel, W. Hochleitner, J. Leitner, M. Haller, A. Geven, M. Tscheligi, "Using heuristics to evaluate the overall user experience of video games and advanced interaction games,". In: *Evaluating user experience in games*. Springer London, 2010, pp. 233-256.

[28] K. Suhonen, H. Väätäjä, "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*. ACM, 2010, pp. 147-150.

[29] J. Paavilainen, "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*. ACM, 2010, pp. 56-65.

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

[31] Y. H. Liao, C. Shen, "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*. IEEE, 2012, pp. 192-196.

[32] T. Hynninen, "First-person shooter controls on touchscreen devices: A heuristic evaluation of three games on the iPod touch" M.Sc. Thesis, Department of Computer Sciences, University of Tampere, Tampere, Finland, 2012.

[33] A. Ponnada, A. Kannan, "Evaluation of mobile games using playability heuristics," In: *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*. ACM, 2012, pp. 244-247.

[34] K. W. Carmody, "Exploring serious game design heuristics: a delphi study," Doctoral dissertation, Northeastern University, 2012.

[35] J. L. G. Sánchez, F. L. G. Vela, F. M. Simarro, N. Padilla-Zea, "Playability: analysing user experience in video games," *Behaviour & Information Technology*, 2012, v31,n.10: pp.1033-1054.

[36] H. Mohamed, A. Jaafar, "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,* 2013.

[37] G. Ulger, "Playability heuristics for mobile games using touchscreen displays," M.Sc. dissertation, Middle East Technical University, 2013.

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

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

[40] F. Rodio, J. M. Bastien, "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*. ACM, 2013, pp. 89.

[41] J. N. Marciano, L. C. Miranda, E. E. C. Miranda, "Evaluating multiple aspects of educational computer games: literature review and case study," *International Journal of Computer Games Technology*, 2014, v.2014, n.14.

[42] O. A. Wodike, G. Sim, M. Horton, "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*. BCS, 2014, pp. 353-358.

[43] S. Soomro, W. F. W. Ahmad, S. Sulaiman, "Evaluation of mobile games with playability heuristic evaluation system," In: *Computer and Information Sciences (ICCOINS), 2014 International Conference on*. IEEE, 2014, pp. 1-6.

[44] M. Hara, S. Ovaska, "Heuristics for motion-based control in games," In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM, 2014, pp. 697-706.

[45] N. Jerzak, F. Rebelo, "Serious games and heuristic evaluation–the cross-comparison of existing heuristic evaluation methods for games," In: *International Conference of Design, User Experience, and Usability*. Springer, Cham, 2014, pp. 453-464.

[46] C. Hochleitner, W. Hochleitner, C. Graf, M. Tscheligi, "A heuristic framework for evaluating user experience in games. In: *Game User Experience Evaluation*. Springer International Publishing, 2015, pp. 187-206.

[47] M. B. Barbosa, A. B. Rêgo, I. De Medeiros, "HEEG: Heuristic Evaluation for Educational Games," In: *Proceedings of SBGames 2015*, 2015, pp. 224-227.

[48] H. Desurvire, C. Wiberg, "User Experience Design for Inexperienced Gamers: GAP—Game Approachability Principles," In: *Game User Experience Evaluation*. Springer International Publishing, 2015, p. 169-186.

[49] Y. R. Guo, D. H. Goh, "Heuristic evaluation of an information literacy game," In: *Digital Libraries: Knowledge, Information, and Data in an Open Access Society*. Springer International Publishing, 2016, pp. 188-199.

[50] M. Zhu, F. Zhao, X. Fang, C. Moser, "Developing Playability Heuristics Based on Nouns and Adjectives from Online Game Reviews," *International Journal of Human–Computer Interaction*, 2017, v.33,n.3, pp. 241-253.

[51] R. M. Gagne, L. J. Briggs, *Principles of instructional design*. Holt, Rinehart & Winston, 1974.

[52] J. M. Keller, "Using The ARCS motivational process in computer□ based instruction and distance education," *New directions for teaching and learning*, 1999, v.1999,n.78, pp. 37-47.

[53] D. A. Norman, *The design of everyday things.* London, MIT Press, 1998.

[54] S. Papaloukas, K. Patriarcheas, M. Xenos, "Usability assessment heuristics in new genre videogames," In: *Informatics, 2009. PCI'09. 13th Panhellenic Conference on*. IEEE, 2009, pp. 202-206.

[55] W. IJsselsteijn, Y. A. W. De Kort, K. Poels, "The game experience questionnaire," *Manuscript in preparation*, 2008

[56] M. Csikszentmihalyi, *The psychology of optimal experience.* New York, 1990.

[57] G. Skulmoski, F. Hartman, J. Krahn, "The Delphi method for graduate research," *Journal of Information Technology Education: Research*, 2007, v.6,n.1, pp. 1-21.

[58] K. Cuhls, "Delphi method," Fraunhofer Institute for Systems and Innovation Research. Germany, 2003.

[59] J. Nielsen, "Severity ratings for usability problems," *Papers and Essays*, 1995, v.54, pp. 1-2.

[60] K. Baker, S. Greenberg, C. Gutwin, "Empirical development of a heuristic evaluation methodology for shared workspace groupware," In: *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. ACM, 2002, pp. 96-105.

[61] F Garzotto, "Investigating the educational effectiveness of multiplayer online games for children," In: *Proceedings of the 6th international conference on Interaction design and children*. ACM, 2007, pp. 29-36.

[62] T. L. Saaty, "How to make a decision: the analytic hierarchy process," *European journal of operational research*, 1990, v.48,n.1, pp. 9-26.