Critical Review on Video Game Evaluation Heuristics: Social Games Perspective

Janne Paavilainen
University of Tampere
Kanslerinrinne 1
33014 Tampereen Yliopisto, Finland
janne.paavilainen@uta.fi

ABSTRACT

This paper presents the first step in creating design and evaluation heuristics for social games which emerge from the domain of social media. Initial high level heuristics for social games are offered by reviewing four existing video game heuristic models and analyzing two social games design frameworks.

Categories and Subject Descriptors

H.5.3 [**Information Interfaces and Presentation**]: User Interfaces – *Evaluation/Methodology.*

General Terms

Design, Human Factors

Keywords

Video Games, Usability, Playability, Heuristics, Evaluation, Social Games

1. INTRODUCTION

A new genre of video games labeled "social games" has emerged from the domain of social media [17, 42]. Social games are web-based browser games played on social network services such as Facebook, MySpace and Twitter. On Facebook, games like *FarmVille*[1] attract over 80 million gamers monthly and there are several other games with over 10 million monthly gamers [1]. Facebook itself has passed the 400 million users milestone [8] and the large user base, easy accessibility, viral distribution and various business models have attracted many developers to create games for social networking services [13].

Usability inspection methods, such as heuristic evaluation, have gained increasing interest among game researchers in the last decade as several video game specific heuristic evaluation models have been published [e.g. 6, 7, 12, 19, 38, 41]. Heuristic evaluation is an expert review method in which a group of experts evaluate a product based on a set of heuristics which are rule of thumb statements reflecting the desirable aspects of a given product [34]. Heuristic evaluation was introduced by Nielsen and Molich in 1990 and it was originally used for evaluating the user interface of productivity software [33]. Heuristics can be also used as design guidelines in the development phase.

Several authors have noted that video games require heuristics of their own [e.g. 6, 12, 19]. In addition to usability, video game heuristics address issues concerning playability. Playability is a rather vague term which, unlike usability, does not have a standardized definition [cf. 16]. Several authors have offered definitions for playability [e.g. 9, 11, 18, 21] but none of them could be considered as a *de facto* standard in the academia.

As there is no commonly agreed definition for playability, there are numerous heuristic models for evaluating playability of video games. To add to the diversity, there are also genre specific heuristics for mobile games [19] and advanced table top games, for example [24]. The authors of video game heuristics claim that their heuristics are useful [e.g. 6, 19, 38] but comparative research to confirm these claims has been limited thus far [21].

Currently, there are no heuristics developed for social games. This is the point of interest in this paper and in forthcoming papers of our research. To start developing social game heuristics, we need to understand how video game heuristics are developed and what the characteristics of social games are. By reviewing video game heuristics and their development methodologies and by analyzing the design frameworks of social games, this paper presents initial high level heuristics for social games along with suggested methodologies for refining them in the future

This paper is structured as follows: the next chapter presents published video game heuristics in chronological order. The third chapter presents a critical comparative review on four different heuristic models and methodology recommendations for creating new heuristics for social games. The fourth chapter presents and analyses two different social game design frameworks and discusses their relation to the reviewed heuristics. The fifth chapter presents initial high level social game heuristics based on the frameworks and on our experience of the domain. The paper ends with discussion, conclusions and ideas for future work.

2. HEURISTICS FOR VIDEO GAMES

The first heuristics concerning video games were presented by Malone in the early 1980s [28, 29]. Malone's heuristics for educational video games feature three categories: Challenge, Fantasy and Curiosity. These three categories contain numerous

ACM Future Play2010, May 6–7, 2010, Vancouver, BC, Canada.

[1] http://www.facebook.com/apps/application.php?id=102452128776

ACM FuturePlay 2010, May 6-7, 2010, Vancouver, BC. Canada

sub-components such as goals, uncertain outcome, intrinsic and extrinsic fantasies and cognitive curiosity.

Clanton's short paper from 1998 [5] presents 15 principles for computer game design for gaining and keeping the player's interest. Clanton suggests that human-computer interaction (HCI) in computer games can be divided into Game Interface, Game Mechanics and Game Play. Clanton's principles can be interpreted to be heuristics, each featuring a title and a short description.

Falstein and Barwood have been collecting rules of game design "*that can be used by designers to make better games*" [10]. Known as The 400 Project, this work can be traced back into 2001 and the list currently contains 112 rules with the latest update dating from March 2006. These rules could also be used as design or evaluation heuristics and they are categorized by their ID number, imperative statement, explanation, domain and contributors.

Federoff's Master's Thesis [12] can be considered as the first modern video game specific heuristic model due its structure and design method (literature review accompanied by empirical study). Based on a literature review, Nielsen's ten heuristics [35] and game designer observations and interviews, Federoff presents a list of 40 heuristics categorized similarly to Clanton's model into Game Interface, Game Mechanics and Game Play.

Desurvire et al. presented Heuristics for Evaluating Playability (HEP) in 2004 [6]. Based on literature reviews and feedback from game designers and playability experts, the model consists of 43 heuristics divided into four categories: Game Play, Game Story, Mechanics and Usability. The model was validated by evaluating an early Flash-game prototype with a standard user study (playtesting) and comparing the results to the HEP evaluation. The results indicated that HEP was able to find more playability problems than the user study.

Koivisto and Korhonen presented playability heuristics for mobile games in 2006 [19]. Heuristics were created by studying mobile use context, literature review and evaluating mobile games. Authors present 29 heuristics in three categories: Game Usability, Mobility and Gameplay[2]. These heuristics were designed for Nokia [36, 37] and later a new category was added with eight additional heuristics for mobile multiplayer games [20]. The heuristics were validated by evaluating single and multiplayer mobile games, and multiplayer PC games.

In 2007, Schaffer presented a white paper introducing heuristics for usability in games [41]. According to Schaffer, earlier heuristics [6, 12] lacked concrete examples, making them less clear for practitioners. Schaffer's heuristics are based on literature and on his own expertise from the field of HCI. The heuristics are divided into five categories: General, Graphical User Interface, Gameplay, Control Mapping and Level Design. These categories contain a total of 21 heuristics with sub-notes and the white paper features screenshot examples of each heuristic.

Pinelle et al. presented usability heuristics for video games in 2008 [38]. The authors present ten heuristics that are focused on game usability which "*does not address issues of entertainment, engagement, and storyline, which are strongly tied to both artistic… and technical issues*". Later, Pinelle et al. introduced 10 additional usability heuristics for multiplayer games [39]. The basis of these heuristics is the analysis of commercial game reviews from the GameSpot[3] and GameSpy[4] websites. The heuristics were validated by evaluating single and multiplayer PC games and the multiplayer heuristics were compared against groupware heuristics in a user study.

In 2009, Desurvire and Wiberg presented a refined list of the original HEP heuristics [7]. The revised list contains 19 heading-level heuristics in three categories: Game Play, Coolness/Entertainment/Humor/Emotional Immersion and Usability & Game Mechanics. The model is based on selected literature as the authors do not cite the more recent works of Korhonen and Koivisto [19], Schaffer [41] and Pinelle et al. [38]. The model is especially intended for action adventure, first-person shooter and real-time strategy games. The revision process was based on discussions with representatives of the game industry and the use of a questionnaire study with 52 valid respondents.

Lastly, Köffel et al. has recently presented a synthesis of the earlier heuristics [25]. The authors handpicked 29 heuristics from various authors and implemented 10 additional heuristics from earlier study concerning advanced table top games [24].

3. REVIEW ON HEURISTICS

This section presents a critical comparative review on four different video game evaluation heuristic models. These models have been presented by four different authors. In two cases, the authors have extended their original work with additional heuristics concerning multiplayer games. The selected models for review are:

1. Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games by Federoff [12]
2. Using Heuristics to Evaluate the Playability of Games by Desurvire et al. [6]

3a. Playability Heuristics for Mobile Games by Korhonen and Koivisto [19]

3b. Playability Heuristics for Mobile Multi-Player Games by Korhonen and Koivisto [20]

4a. Heuristic Evaluation for Games: Usability Principles for Video Game Design by Pinelle et al. [38]

4b. Usability Heuristics for Networked Multi-Player Games by Pinelle et al. [39]

These models were selected because they are considered to be comparable due the similarities in their format and according to their authors; they are useful in evaluating video games. Each model has been developed in a different way and there is variance in the heuristics and their structure. Some more recent models [e.g. 41, 7] were left out of this review due to the limitations of this paper and because their development methods do not bring anything new onto the table.

The review in the following sub-chapters is based on the specifications of the heuristics, i.e. development, validation and content. Development presents how the heuristics were created,

[2] Korhonen and Koivisto use the term "Gameplay" in place of "Game play".

[3] http://www.gamespot.com

[4] http://www.gamespy.com

validation presents if and how the heuristics were validated and content focuses on the type, wording, abstraction level, number and structuring of the heuristics. As this paper focuses on social games, references to social aspects in the heuristics are also examined. The last sub-chapter summarizes and compares the findings, and presents recommendations concerning the methodology for creating new heuristics for social games.

3.1 Federoff (2002)

Federoff's heuristics are presented in her master's thesis from 2002 [12]. The heuristics were created by reviewing usability literature and executing a field study in a game development company. The goal of the study was to examine the implicit and explicit heuristics used in game design, and the evaluation measures used in a game development company.

The development process for this model is questionable. The methodology resembles that of a case study, which does not produce results that can be generalized. This has been noted by other authors as well [25, 44]. Federoff interviewed and observed only one development team with five members, which might make the heuristics biased. To add validity to this method, more designers from different teams who work on different game development projects should be interviewed.

The main critique with this model is that it has not been validated in any way. This was also pointed out by the author herself. Validation was left out for future work as it was not considered to be in the scope of the thesis. Due to the lack of validation, many of Federoff's heuristics are challenging to use in an evaluation. For example, the heuristics "*Play should be fair*" and "*Create a great storyline*" are too abstract. The heuristic "*A good game should be easy to learn and hard to master*" (a quote from Nolan Bushnell) is vague and hard to be evaluated in a game. On the level of description, the heuristic: "*If the game cannot be modeless, it should feel modeless to the player*" is hard to understand without consulting the actual thesis for explanation. The heuristics "*Game play should be balanced so that there is no definite way to win*" and "*There must not be any single optimal winning strategy*" are very similar to each other, creating redundancy. In addition to the critique from Wiberg et al. [44] and Köffel et al. [25], Schafer [41] considers that Federoff's heuristics are relatively vague and difficult to be implemented during the design process.

The number of heuristics is high, reaching a total of 40. The high number of heuristics was considered problematic in an earlier comparison study with models containing 43 and 29 heuristics [21]. Although the number of heuristics is high, none of them refer to the social aspects of video games. The category of Game Mechanics contains only two heuristics and one shared heuristic with the category of Game Play, which questions the need for a separate category for Game Mechanics. The category of Game Interface contains heuristics which work on very different levels of abstraction. For example, "*For PC games, consider hiding the main computer interface during game play*" is a platform specific technical heuristic where as "*Do not expect the user to read a manual*" could be a design guideline for almost any software or hardware item like a DVD player, for example.

3.2 Desurvire et al. (2004)

Heuristics for Evaluating Playability (HEP) was presented by Desurvire et al. in the CHI'04 conference as an extended abstract in the section of late breaking results papers [6]. The authors noted that earlier video game heuristics had not been verified in any way, nor was there a comprehensive list compiled. This was the rationale to create a new heuristic model.

The model is based on a literature review. It was also reviewed by several playability experts and game designers. Validation was done by evaluating an early phase Flash game prototype and comparing the HEP evaluation results against the results of a standard user study. The results showed that a HEP evaluation was able to find more playability problems than a user study.

Due to the short nature of the extended abstract, a lot of important information about the development of the model was not presented. For example, the authors state that the model was reviewed by "*several playability experts and game designers*" but their number or expertise is not revealed or explained. Furthermore, there is no information regarding whether the heuristics were modified after the review or which heuristics were derived from what source. Some of the heuristics are same than in Federoff's model.

The validation procedure of the heuristics is questionable as there was only one evaluator using the HEP model and only four testees in the user study. Only one early phase prototype game was evaluated which did not allow any gameplay. Because of these issues in the validation, the author's statement that HEP is "*very useful in creating highly usable and playable game design*" is dubious, especially when gameplay was not possible with the evaluated game.

Due to the shortfalls of the validation procedure, there are many problematic heuristics in the model. The lack of explanations for the heuristics underlines these problems. The problems have also been discussed elsewhere [21] and the model is also criticized by Schaffer, who considers that the heuristics are relatively vague and difficult to be implemented in the game design process [41]. Wiberg et al. consider that HEP should be modified and validated for each new game [44].

Some heuristics are not suitable for heuristic evaluation at all. The heuristic "*The game is enjoyable to replay*" is a subjective issue which cannot be evaluated by an expert review method during the development process. The authors also noted this and agree in the paper that it is an issue which can only be found out by observing playtesting. For one reason or another, the authors included this heuristic in the HEP model in any case.

Further examples of subjective heuristics are "*The player spends time thinking about possible story outcomes*" and "*Player is interested in the character because (1) they are like me, (2) they are interesting to me, (3) the characters develop as action occurs*". Heuristics like "*The game is fun for the player first, the designer second and the computer third. That is, if the non-expert player's experience isn't put first, excellent game mechanics and graphics programming triumphs are meaningless*" are difficult to be used in evaluation [21]. When compared to "*The first player action is painfully obvious and should result in immediate positive feedback*", it is clear that the heuristics lack consistency on the level of abstraction and detail. Many of the heuristics seem to be in the form of design requirements and they have not been written down with evaluation in mind. Further examples of such heuristics are "*Players discover the story as part of game play*", "*Player is taught skills early that you expect the players to use later, or right before the new skill is needed*" and "*Pace the game*

to apply pressure but not frustrate the player. Vary the difficulty level so that the player has greater challenges as they develop mastery. Easy to learn, hard to master".

The number of heuristics is even higher than in Federoff's model. A total of 43 heuristics are presented but none of them refer to social aspects in games. An earlier study [21] revealed that some heuristics seem to be in the wrong category. For example, "*Players discover the story as part of game play*" is in the Game Play category while "*Player experiences fairness of outcomes*" is in the Game Story category. By looking at the wording, it could be interpreted that the categorization of these heuristics should be vice versa. On the other hand, the one in Game Story category could mean that the player experiences the fairness of outcome in relation to the game story, but due to a lack of explanation for the heuristic, it remains unclear.

3.3 Korhonen and Koivisto (2006, 2007)

Korhonen and Koivisto presented a heuristic model for evaluating mobile games in 2006 [19]. The model was introduced in the MobileCHI'06 conference as a full paper. According to the authors, earlier models did not cater to mobile games, they lacked applicability and were ambiguous. Later the original model was extended with mobile multiplayer heuristics, which were presented in a full paper in the DIMEA'07 conference [20]. According to the authors, the model is modular in a sense that the evaluator can use the heuristic categories independently.

The first model was based on a literature review, a mobile use context study, and an evaluation of three mobile games. Validation was done by evaluating five additional mobile games. The additional heuristics for multiplayer games were based on a literature review and an evaluation of three mobile multiplayer games. These heuristics were validated informally by evaluating six PC multiplayer games. All evaluations were done with a heuristic evaluation method. According to the authors, the heuristics that were created are useful in evaluating single and multiplayer mobile games. Due the modular nature of the heuristics, they can also be used for evaluating games for other platforms, as well [21].

The development method relies heavily on game evaluations. These heuristics have not been compared against user studies, which would further validate the claims of their usefulness. However, due to the rigorous testing of the heuristics, most of them are easy to understand by their heading level descriptions. The heuristics are also consistent in their level of abstraction, making them feasible for evaluation purposes. There is also an additional white paper and a web resource available for explaining the heuristics thoroughly [36, 37].

In an earlier study by Korhonen et al. [21], the evaluators considered that there are heuristics that overlap in their heading level descriptions. Heuristics "*There are no repetitive or boring tasks*" and "*The game does not stagnate*" looked similar to some evaluators.

There are also some heuristics that could possibly be merged. Examples of these are Game Usability heuristics "*Control keys are consistent and follow standard conventions*" and "*Game controls are convenient and flexible*". Another example is the Multiplayer heuristics "*The game supports communication*" and "*There are reasons to communicate*".

The total number of heuristics is high, 37 in total, which is slightly less than in Federoff's and Desurvire et al.'s models. One Gameplay heuristic is related to social aspects, "*The player sees the progress in the game and can compare the results*". Most of the multiplayer heuristics are related to social aspects, like "*The game helps the player to find other players and game instances*" and "*The game supports groups and communities*". There is one exception, "*The design hides the effects of network*" does not have a social aspect as it addresses a technical network problem.

3.4 Pinelle et al. (2008, 2009)

Pinelle et al. presented usability heuristics for video games in a full paper section in the CHI'08 conference [38]. The authors stated that earlier heuristics presented by Clanton, Federoff and Desurvire et al. were strongly focused on engagement and fun, and did not cover usability in detail. Although Korhonen and Koivisto had created usability heuristics for games earlier, their study was not acknowledged. Later, Pinelle et al. presented usability heuristics for multiplayer games. These heuristics were introduced in a full paper in the GROUP'09 conference [39].

The first set of heuristics was created by three evaluators who reviewed 108 commercial game reviews from the GameSpot website. These heuristics were validated by evaluating a demo version of a PC game. The second set of heuristics was created in a similar manner by reviewing a total of 382 reviews from the GameSpot and GameSpy websites. The multiplayer heuristics were validated by having two groups evaluate two multiplayer video games. One group used the newly designed heuristics and the other used groupware heuristics by Baker et al. [2]. The group that used the newly designed heuristics was considered more effective in finding multiplayer usability problems. According to the authors, the heuristics created are useful in the evaluation of usability for video games.

The heuristics of Pinelle et al. are not without problems. The authors note that the models only address usability issues so evaluators need to consider other models for evaluating engagement and fun. However, the larger debate is the development method for the heuristics. Using game reviews as a data source for usability problems is interesting but it certainly has its flaws. Journalists are not usability evaluators and they do not point out every usability problem found. The goals and the target audience for a game journalist and usability evaluator are different.

The result of this approach is a low usability-problem-per-game ratio. The first study had an average of 2.64 usability problems per game, which is significantly lower compared to the findings of other authors who have used heuristic evaluation for evaluating video games [e.g. 6, 19, 21]. For the second study, 213 reviews of a total of 382 discussed problems relating to the multiuser aspects of the game. Out of the 213 reviews, the authors identified 147 multiplayer usability problems (0.69 problems per game) and 187 related problems (0.87 problems per game) making these numbers even lower than in the first study.

Köffel et al. [25] consider that the data for the first study might be biased, as the reviews were taken from GameSpot only, which was under heavy criticism in November 2007 due to "*allegedly firing a reviewer due to the negative reviews of a game published*

by a financial sponsor of GameSpot"[5]. Game journalism in general has also received criticism for being on the leash of the game industry [32].

Proper validation for the heuristics is lacking. The first set of heuristics was validated by evaluating only one PC game demo with five evaluators. In validating the multiplayer heuristics, the authors used groupware heuristics from Baker et al. [2] for comparison against the newly designed heuristics. The authors' finding that the multiplayer usability heuristics were better than the groupware heuristics in finding usability problems in multiplayer games is axiomatic. Comparison with the multiplayer heuristics from Korhonen and Koivisto would have been more appropriate. The comparison study featured rather small groups, five persons in each, which results in low statistical power and suffers from the wildcard effect as reported by Gray and Salzman [14]. The heuristics of Pinelle et al. have not been compared with playtesting.

Despite the unorthodox development method and shortfalls in the validation, the defined heuristics are presented clearly with explanations. Unlike previous authors, Pinelle et al. present the heuristics in similar fashion to Nielsen [35], i.e. each heuristic has a heading with a one paragraph description, all presented in a single page. There are no overlapping heuristics and no additional documents are needed for understanding them. The number of heuristics is significantly lower than in other models, but the heuristics only cover usability issues in single and multiplayer games. The descriptions are detailed and cover many aspects under the heading of the heuristic.

Social aspects are found in nine heuristics in the multiplayer usability category. The heuristics are similar to the ones of Korhonen and Koivisto, addressing issues like session management, communication, social interaction, etc. The only multiplayer usability heuristic that was not considered to have a social aspect was "*Reduce game-based delays*". This is a rather similarly technically orientated heuristic to Korhonen and Koivisto's, "*The design hides the effects of network*".

3.5 Summary and Recommendations

The reviewed models had significant differences in their development and validation methods, as well as in content. Due to the limited scope of this paper, a side-by-side comparison between each heuristic is neither possible nor necessary.

Considering development methods, the approach by Korhonen and Koivisto seems to be most valid. By understanding the use context and evaluating several games from the studied domain, it can be expected that the heuristics have value in the sense of consistency and generalizability. The method by Pinelle et al. can be seen to be problematic, due to the use of second hand data and the question here is if the end justifies the means. Federoff and Desurvire et al. base their heuristics on theory only and their work resembles case studies with low power of generalization.

All models are lacking in validation. Korhonen and Koivisto validated their heuristics with only heuristic evaluations without comparing the results to other methods like playtesting. Pinelle et al. compared their multiplayer usability heuristics against groupware heuristics, but the statistical power of the comparison is low and it suffers from the wildcard effect [14]. The study by Desurvire et al. suffers from the same problem and Federoff's heuristics were not validated at all.

For the clarity and consistency of the heuristics, the models from Korhonen and Koivisto and from Pinelle et al. succeed well. Both of them contain clearly presented heuristics, which are explained thoroughly. The model by Korhonen and Koivisto relies on external documents, whereas Pinelle et al. have integrated everything in a single piece of paper. The heuristics of Federoff and Desurvire et al. lack in clarity and consistency, and their heuristics have also been criticized by other researchers.

The coverage of heuristics varied. Whereas Pinelle et al. only focus on single and multiplayer usability issues Korhonen and Koivisto present a larger scope of categories for evaluating playability issues. Federoff and Desurvire et al. also cover both usability and playability issues but without multiplayer heuristics.

The number of heuristics is quite high in each model, even in the model of Pinelle et al., when considering that it only covers usability issues. The high number of heuristics could be explained by the complexity of modern video games in general. Especially genre breaking games such as *Grand Theft Auto IV*[6] and *Fallout 3*[7] contain aspects of many genres, making them complex and thus enabling the use of a large number of heuristics.

An interesting aspect is that although games are inherently social [e.g. 17, 42], only one non-multiplayer heuristic contained a social aspect. It might be that social aspects in games have not been considered outside of multiplayer games.

When it comes to recommendations for the development of social game heuristics, we consider that the methodology used by Korhonen and Koivisto to be most appropriate. Understanding the domain thoroughly and developing and validating the heuristics iteratively through extensive game evaluations seems to lead to consistent heuristics. However, we consider that the new heuristics need to be evaluated against other methods like playtesting to increase their validity. Special attention should be given to the threat categories described by Gray and Salzman when arranging and executing comparison studies for heuristic evaluation models [14]. Interviewing game designers and active gamers can be used as supplementary research in the process.

4. SOCIAL GAMES PERSPECTIVE

As social games are a new phenomenon, there has been little in the way of academic research delving into them. This section presents two social games design frameworks. These frameworks are used to understand the domain of social games. The frameworks are presented, analyzed and finally reflected against the reviewed heuristics.

4.1 Social Games Design Frameworks

Järvinen has presented five design drivers for social games [17]. His framework is based on the earlier work of Rao, who identified three factors that characterize the playful use of Facebook: Physicality, Spontaneity and Inherent Sociability [40]. For his framework, Järvinen has added two additional factors: Narrativity and Asynchronicity as discussed by Bogost [3].

[5] See http://en.wikipedia.org/wiki/GameSpot#Gerstmann_dismissal for related discussion.

[6] http://www.rockstargames.com/IV/

[7] http://fallout.bethsoft.com/index.html

According to Järvinen, Spontaneity is closely related to accessibility and familiarity. Seemingly complex tasks (like fighting a gangster in *Mafia Wars*[8]) are executed with single clicks of mouse and familiar topics (like gardening in *FarmVille*) are easy to understand and access. Symbolic Physicality manifests in actions like virtual poking, beer sharing, high-fiving, sending gifts, etc. These actions try to bring physicality into a non-physical social space. Inherent Sociability is obvious in social networks as, more often than not, social games feature in-game attachments to the player's social contacts in her network. These contacts can have interactive or non-interactive roles in the gameplay. Narrativity is found in the stylized narrative rhetoric which is apparent in many social games. Rather than just presenting the explicit gameplay results to the player, these vivid stories are broadcasted through the player's news feed and add depth to the gameplay. Narratives that are posted to the news feed also act as links for others to join in the game. Asynchronicity manifests in the interaction of players and the game worlds. The player can quit and resume the game at any time, always being able to continue on from where she left off. Games like *FarmVille* are built around asynchronous interaction as the players are expected to check-in from time to time for the harvest. Järvinen has analyzed five social games with these design drivers and provides a set of 12 examples from design patterns, which are pragmatic design suggestions connected to the design drivers [17].

The second framework is presented by Ventrice [43], who states that social games can be defined by three implicit objectives: (1) build a persistent society, (2) maintain a consistent sense of discovery, and (3) spread the game virally.

Persistent society means two things. Persistence refers to the asynchronous nature of social games. They are always on and the players are able to continue on where they have left off. Society refers to cooperation and teamwork with others. Maintaining a consistent sense of discovery means that there is always something new for the player to acquire and experience. Collecting virtual items is a strong motivator for players. The player's ability to express themselves creates an emergent world of new experiences as the player can witness what others have created or discover their own inner muse. Virality comes in two forms, direct and indirect. Direct virality means that the player sends a request or gift for her friend, which acts as an invitation to the game. Indirect virality means that the game broadcasts in-game events into the player's news feed, where it reaches the player's social network [43].

4.2 Analysis of Frameworks

Järvinen and Ventrice have different approaches but their frameworks also have similarities. Järvinen has a broader view, considering play in social network (Facebook) whereas Ventrice is more focused on the exact features of a social game.

Asynchronicity is apparent in both frameworks as well as sociability, where Ventrice promotes design features such as cooperation as a mandatory element for progress in the game. Järvinen also has pragmatic suggestions in the form of design patterns. Some of the design patterns overlap with Ventrice's model, for example:"*persuasive message for inviting into play*", which is discussed by Ventrice in the form of in-direct virality (call-to-action principle). Ventrice dissects virality into two (direct and in-direct) whereas Järvinen embeds viral qualities more pervasively into the whole framework.

Ventrice emphasizes discovery and expression, aspects that are not featured in Järvinen's framework. Järvinen considers spontaneity as an important quality, which is not noted by Ventrice at all. Symbolic physicality is also only covered in Järvinen's framework; however, apart from the obvious "send energy pack" actions in *Mafia Wars,* the exact role of this design driver is a bit vague. References to symbolic physicality in the design patterns include the aforementioned physical contacts (virtual high fives etc.) and friends acting as game commodities. Considering the latter, symbolic physicality as a term might be a bit misleading.

Although the frameworks are not directly comparable, analyzing the frameworks helps to understand the domain of social games and therefore help in the process of developing initial heuristics. This analysis gives a starting point for creating the initial list of high level design and evaluation heuristics for social games.

4.3 Relation to Reviewed Heuristics

When comparing at the reviewed heuristics that contained social aspects and the analyzed frameworks, it becomes clear that the domain of social games has many distinctive features which are not apparent in the current video games heuristics. This is understandable as these heuristics were not created with social games in mind.

Although the multiplayer heuristics from both Korhonen and Koivisto and Pinelle et al. feature social aspects like communication and communities, issues like virality and narrativity are not discussed at all. When examining the multiplayer heuristics more closely as potentially applicable to social games, it becomes apparent they have been developed for more complex games than those on social networks. References to multiple text and voice communication channels or to solutions for managing deviant behavior are not issues in social games, at least not yet. Some heuristics are a bit more relevant to social games, such as "*Support social interaction*" from Pinelle et al. Heuristics like "*The game helps the player to find other players and game instances*" of Korhonen and Koivisto are irrelevant to current social games, as these games automatically take advantage of the player's social network .

The non-social heuristics of the reviewed models remain relevant, however. Social games also require good user interface design which is covered well by the heuristics of Korhonen and Koivisto and Pinelle et al. Gameplay heuristics like "*The first time experience is encouraging*" and "*The game provides clear goals or supports player created goals*" of Korhonen and Koivisto are relevant to all games, regardless of platform. Interestingly, some popular social games like *FarmVille* and *Mafia Wars* actually break traditional heuristics like "*There are no repetitive or boring tasks*" of Korhonen and Koivisto. It might be that the mindset of playing social games is similar to playing casual games, and repetitive, simple tasks are favored for the purposes of relaxation [22]. After all, when looking at the design of social games and casual games side-by-side, the only real difference is the social network and viral distribution channels. It has been noted by industry experts that social games might be eating the markets of casual games [30].

[8] http://www.facebook.com/apps/application.php?id=10979261223

5. INITIAL HEURISTICS

Based on the frameworks by Järvinen and Ventrice, we present initial high level social games heuristics for design and evaluation. These heuristics can be considered as a starting hypothesis, or as a springboard for further research and discussions. Like Järvinen, we have also added new elements based on our ongoing research in the field of social games and heuristic evaluation methods. The initial heuristics are presented in Table 1 below.

Table 1. Initial High Level Heuristics for Social Games.

ID	Description
SG1	**Spontaneity**. Provide easy and quick access to the game as the threshold for play should be as minimal as possible. Use common and familiar themes from popular culture which can be understood easily.
SG2	**Interruptability**. Use game mechanics which support playing in short, sporadic bursts. Use interruptions as an advantage in the design.
SG3	**Continuity**. Provide an asynchronous persistent game world and mechanics that allow the player to feel progress. Provide multi-level reward structures that make the player feel accomplishment in every play session and reward players for coming back into the game.
SG4	**Discovery**. Provide players with new, evolving content and offer an emergent game world. Provide achievements and trophies for players to acquire.
SG5	**Virality**. Use versatile means for direct and in-direct virality. Use "call to action" principle i.e. persuasive, inviting messages, for getting the attention of new players. Provide bonuses that encourage players to send requests and gifts which act as links for new players to start the game.
SG6	**Narrativity**. Use vivid stylized narratives for describing in-game events and broadcast these narratives to engage players and elicit curiosity among others.
SG7	**Sharing**. Provide means for players to share information and in-game resources.
SG8	**Expression.** Allow players to express themselves in the game world. Provide means for expressing game experiences through screenshots and video clips.
SG9	**Sociability**. Use social contacts as assets in the game and make them part of the game mechanics. Support group forming and provide bonuses for communicating and cooperating with contacts.
SG10	**Ranking**. Provide high-score lists for competing with friends. Provide hints and tips on how to climb the ladder and provide reasons for doing so.

We have directly derived Spontaneity and Narrativity heuristics from Järvinen's framework. Continuity reflects the asynchronicity and persistence as presented in both frameworks. Discovery and Expression are directly taken from Ventrice's framework. Virality and Sharing are present (in different levels) in both frameworks and they are major features in social games in general, hence they are presented here as their own heuristics. Sociability is also apparent in both frameworks and here the emphasis is on merging game mechanics with the player's social network. We have also added two heuristics which are not explicitly discussed in the analyzed frameworks.

Firstly, we consider interruptability as an important factor which underlines the sporadic nature and the fragmented rhythm of gameplay in social games. Whereas Spontaneity focuses on how to make the game easily and quickly accessible, Interruptability focuses on how well the game supports short bursts and sudden stops in the gameplay. Secondly, we promote the role of ranking and playful competition in general, as the availability of social networks makes it possible to offer customized high-score lists based on the player's contacts.

It is difficult to draw exact boundaries between the heuristics. For example, spontaneity is connected to interruptability, narrativity is related to virality and expression is one form of sharing. However, we consider that these ten heuristics act as a useful starting point for further studies and they will be defined more accurately in the future as our research continues.

6. DISCUSSION

Based on the reviewed heuristics, we can see a gradual improvement over time in the validity of development methodologies. Validation is an important part of the process as it seems to affect the clarity and consistency of the heuristics. There is clearly a need for more rigorous comparison studies based on experiments with users. Gray and Salzman have noted serious deficiencies in these kinds of studies in the domain of productivity software heuristics and they have identified threat categories which hinder the validity of comparison experiments [14]. These threats are visible in the validation experiments of the current video game heuristics also, which suffer from wildcard effects and low statistical power.

The reviewed models contain a high number of heuristics, from 20 to 43. The high number of heuristics was considered a problem in an earlier study [21] and when considering Miller's golden rule of 7±2 [31], it is clear that the cognitive load becomes high for a user, as well. When presenting the first version of usability heuristics for user interfaces, Nielsen and Molich stated that "*The nine heuristics can be presented in a single lecture and explain a very large proportion of the problems one observes in user interface designs*" [33]. Ling and Salvendy have also suggested that domain specific models should not contain too many heuristics [27]. However, video games are more than just user interfaces. Thus, the high number of heuristics might be justified but, in general, less might be more. This has also been the intention with our heuristics.

The presentation of the heuristics models has been rather uninspired. While Schaffer has used visual aids in the form of screenshots for giving practical examples of the heuristics [41], other authors have settled for simple textual lists. In addition to the high number of heuristics, the ascetic presentation format may discourage active engagement of the heuristics model by users, as navigating through the list becomes tedious. Presenting the heuristics in the form of colourful cards with inspirational figures, for example, could bring playful tangibility to the use. Evaluators considered that this is a good idea in the earlier study [21].

The new interesting genre of social games presents challenges for the previous heuristic models. Although some heuristics are certainly applicable, issues like interruptability, virality and narrativity are poorly, if at all, covered in them. Therefore, it is clear that there is a need for heuristics specifically developed for social games, which would be based on understanding the use

context and characteristics of social games. Interestingly, some of these heuristics can be found as design guidelines and supporting patterns from the earlier research on massively multiplayer mobile phone games [15]. Holopainen et al. present design guidelines such as "*Support short, spontaneous play sessions*" and "*Viral invitations and player recruitment*" along with supporting design patterns like "*interruptability*". It might be that these different design fields might have something to offer to one another.

Currently, the initial high level heuristics serve as inspiration for design and evaluation. They can also be seen as genre description and used for examining contemporary social games. Based on how well the examined social game reflects the heuristic in question, the evaluator could assign a rating based on a scale of one to five, for example. Ventrice has used his design framework in such a way to examine popular social games [43]. However, the problem with such rating is their subjective nature: what is the actual difference between the grades of two and three, for example?

The high level heuristics could be included with sub-level components such as those used in other models [e.g. 7, 28]. For example, we can examine Spontaneity from several different perspectives and the sub-components could be inspiring and absorbing tutorial (gameplay perspective), straight-forward user interface with minimum clicks required for most common actions (UI perspective) or short loading times of the game (technical perspective). The sub-components would offer practicality in the form of lower abstraction level, but the risk of implementing too much information is apparent, making the heuristics tedious to use.

One intriguing approach would be to tag the heuristics with business oriented factors like acquisition, retention, monetization and following i.e. what are the heuristics that support acquiring the players, keeping the players, turning the players into customers and ultimately making them brand loyal. The heuristics could be also merged with our game ideation tools [23] for innovating new types of social games. The aforementioned presentation format of tangible heuristic cards would allow a smooth transition for that domain.

It must be remembered that heuristics are always products of their time. As gaming culture evolves and new genres appear, the heuristics need to be revised accordingly. When considering Facebook for example, even small changes in the application policies might change the way social games are designed [26]. This is a challenge when trying to pin-point the important qualities of today's social games as these may change tomorrow.

Lastly, although the presented frameworks and the initial heuristics for social games try to identify desirable qualities of social games, good games are always more than sum of their features. Like Järvinen notes, implementing each feature might not be enough [17] and on the other hand, each game does not need to reflect all heuristics in its design. Another thing is that social games as a genre is still in its infancy, and there is no explicit definition on what they should be. The gameplay in social games is currently quite simple and minimalistic but that might change in the future as the market matures [4].

7. CONCLUSIONS

We have presented a critical comparative review of four different video game heuristic evaluation models in order to understand development practices and to identify possible heuristics applicable to social games. Based on our review, the methodology for creating the heuristics used by Korhonen and Koivisto seems the most valid, especially when merged with comparative playtesting experiments. We analyzed social games design frameworks to understand the domain and reflected on the reviewed heuristics on these frameworks. It became evident that the current video game heuristics did not align well with the specific needs of social games. Based on the analyzed frameworks and our ongoing research in social games and heuristic evaluation methods, we have presented ten initial high level heuristics for the design and evaluation of social games.

8. FUTURE WORK

More thorough examinations and empirical work with users are needed to evaluate different video game heuristics with the possible threat categories presented by Gray and Salzman kept in mind. Currently our research is focused on interviewing active Facebook users, arranging design workshops with game designers, and analyzing social games with the initial heuristics. The next step is validating the presented heuristics by evaluating social games and comparing the results against playtesting. The presented heuristics are one part in the *SoPlay*[9] project for understanding social playability.

REFERENCES

[1] AppData. 2010. Application Leaderboard for Facebook. Online: http://www.appdata.com/

[2] Baker, K. Greenberg, S. and Gutwin, C. 2002. Empirical Development of a Heuristic Evaluation Methodology for Shared Workspace Groupware. Proc. of the ACM Conference on Computer-Supported Cooperative Work, 96-105.

[3] Bogost, I. 2004. Asynchronous Multiplay: Futures for Casual Multiplayer Experience. Other Player Conference, IT University of Copenhagen, Denmark.

[4] Carless, S. 2009. GDC Austin: Zynga's Reynolds On Social Game Design's Evolution. Gamasutra News. Online: http://www.gamasutra.com/php-bin/news_index.php?story=25332

[5] Clanton, C. An Interpreted Demonstration of Computer Game Design. Proc. of of CHI 1998, 1-2. DOI= http://doi.acm.org/10.1145/286498.286499

[6] Desurvire, H., Caplan, M. and Toth, J.A. 2004. Using Heuristics to Evaluate the Playability of Games. Proc. of ACM SIGCHI '04 Extended Abstracts, 1509-1512. DOI=http://doi.acm.org/10.1145/985921.986102

[7] Desurvire, H. and Wiberg, C. 2009. Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration. Lecture Notes in Computer Science, 5621/2009, 557-556. Online: http://www.springerlink.com/content/cl1w17lp067k39q1/

[8] Dybwad, B. 2010. Facebook Passes the 400 Million User Mark. Mashable. Online: http://mashable.com/2010/02/04/facebook-400-million/

[9] http://soplayproject.wordpress.com

[9] Egenfield-Nielsen, S., Smith, J.H. and Tosca, S.P. 2008. Understanding Video Games: The Essential Introduction. Routledge.

[10] Falstein, N. and Barwood, H. 2006. The 400 Project. Online: http://theinspiracy.com/400_project.htm.

[11] Fabricatore, C., Nussbaum, M., Rosas, R. 2002. Playability in Action Videogames: A Qualitative Design Model. Human-Computer Interaction 17, 311-368. DOI= http://dx.doi.org/10.1207/S15327051HCI1704_1

[12] Federoff, M.A. 2002. Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games. Master's Thesis. Department of Telecommunications, Master of Science. Indiana University, Indiana.

[13] GP Bullhound. 2010. Social Gaming: The Fastest Growing Segment in the Games Market. Online: http://www.gpbullhound.com/researchpdf/GPB%20Social%20Gaming%20Report%20-%20March%202010%20%283%29.pdf

[14] Gray, W. D. and Salzman, M. 1998. Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. Human-Computer Interaction 13 (3), 203-261.

[15] Holopainen, J., Korhonen, H., Ollila, E., Nenonen, V., Björk, S., Peitz, J. and Davidsson, O. 2008. Design Kit: Massively Multiplayer Mobile Phone Games. IPerG Deliverable D5.10. Online: http://iperg.sics.se/Deliverables/D5.10.pdf

[16] ISO. 1998. ISO 9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on usability. ISO International Standard.

[17] Järvinen, A. 2009. Game Design for Social Networks. Interaction Design for Playful Dispositions. Proc. of the 2009 ACM SIGGRAPH Symposium on Video Games. DOI= http://doi.acm.org/10.1145/1581073.1581088

[18] Järvinen, A., Heliö, S., Mäyrä, F. 2002. Communication and Community in Digital Entertainment Services. University of Tampere.

[19] Korhonen, H. and Koivisto, E.M.I. 2006. Playability Heuristics for Mobile Games. Proc. of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services, 9-16. DOI= http://doi.acm.org/10.1145/1152215.1152218

[20] Korhonen, H. and Koivisto, E. M.I. 2007 Playability Heuristics for Mobile Multi-Player Games. Proc. of the 2nd International Conference on Digital Interactive Media in Entertainment and Arts, 28-35. DOI= http://doi.acm.org/10.1145/1306813.1306828

[21] Korhonen, H., Paavilainen, J. and Saarenpää, H. 2009. Expert Review Method in Game Evaluations – Comparison of Two Playability Heuristic Sets. Proc. of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era, 74-81. DOI= http://doi.acm.org/10.1145/1621841.1621856

[22] Kuittinen, J., Kultima, A., Niemelä, J. and Paavilainen, J. 2007. Casual Games Discussion. Proc. of the 2007 conference on Future Play, 105-112. DOI= http://doi.acm.org/10.1145/1328202.1328221

[23] Kultima, A., Niemelä, J., Paavilainen, J. and Saarenpää, H. 2008. Designing Game Idea Generation Games. Proc. of the 2008 conference for Future Play, 137-144. DOI= http://doi.acm.org/10.1145/1496984.1497007

[24] Köffel, C. and Haller, M. 2008. Heuristics for the Evaluation of Tabletop Games. Evaluating User Experiences in Games, Workshop at the 2008 Conference on Human Factors in Computing Systems.

[25] Köffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A. and Tscheligi, M. 2010. Using Heuristics to Evaluate the Overall User Experience of Video Games and Advanced Interaction Games. In Bernhaupt, R (Ed.) Evaluating User Experience in Games. Springer.

[26] Laakkonen, J. 2009. Facebook Policy Changes Are a Tectonic Shift for Social Game Developers. Sombiz. Online: http://www.sombiz.net/content/facebook-policy-changes-are-tectonic-shift-social-game-developers

[27] Ling, C. and Salvendy, G. 2005. Extension of Heuristic Evaluation Method: A Review and Reappraisal. Ergonomia 27 (3), 179-197.

[28] Malone, T.W. 1980. What Makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games. Proc. of the 3rd ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems, 162-169.

[29] Malone, T.W. 1982. Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games. Proc. of ACM SIGCHI, 63-68. DOI= http://doi.acm.org/10.1145/800049.801756

[30] Mayer, A. 2009. Ten Reasons that Social Games May Eat the Casual Market. Inside Social Games. Online: http://www.insidesocialgames.com/2009/02/10/ten-reasons-that-social-games-may-eat-the-casual-market/

[31] Miller, G. A. 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review 63. 81-97.

[32] Nieborg, D. B. and Sihvonen, T. 2009. The New Gatekeepers: The Occupational Ideology of Game Journalism. Proc. of DiGRA 2009. Online: http://www.digra.org:8080/Plone/dl/db/09287.29284.pdf

[33] Nielsen, J. and Molich, R. 1990. Heuristic Evaluation of User Interfaces. Proc. of ACM SIGCHI, 249-256. DOI= http://doi.acm.org/10.1145/97243.97281

[34] Nielsen, J. 1994. Heuristic Evaluation. In Nielsen, J. and Mark, R.L. (Eds.) Usability Inspection Methods. New York, NY, USA: John Wiley & Sons Ltd.

[35] Nielsen, J. 2005. Ten Usability Heuristics. Useit.com. Online: http://www.useit.com/papers/heuristic/heuristic_list.html

[36] Nokia. 2006. Mobile Game Playability Heuristics. Online: http://bit.ly/nokia_heuristics_pdf

[37] Nokia. 2010. Mobile Game Playability Heuristics. Forum.Nokia.com. Online: http://bit.ly/nokia_heuristics

[38] Pinelle, D., Wong, N., Stach, T. 2008. Heuristic Evaluation for Games: Usability Principles for Video Game Design. Proc. of ACM SIGCHI, 1453-1462. DOI= http://doi.acm.org/10.1145/1357054.1357282

[39] Pinelle, D., Wong, N., Stach, T. and Gutwin, C. 2009. Usability Heuristics for Networked Multiplayer Games. Proc. of the ACM 2009 International Conference on Supporting Group Work, 169-178. DOI= http://doi.acm.org/10.1145/1531674.1531700

[40] Rao, V. 2008. Facebook Applications and Playful Mood: The Construction of Facebook as "Third Place". Proc. of the 12th international conference on Entertainment and media in the ubiquitous era, 8-12. DOI= http://doi.acm.org/10.1145/1457199.1457202

[41] Schaffer, N. 2007. Heuristics for Usability in Games. White Paper. Online: http://www.playerfriendly.com/files/heuristics.pdf

[42] Stenros, J., Paavilainen, J. and Mäyrä, F. 2009. The Many Faces of Sociability and Social Play in Games. Proc. of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era, 82-89. DOI= http://doi.acm.org/10.1145/1621841.1621857

[43] Ventrice, T. 2009. Building the Foundation of a Social Future. Gamesutra Feature. Online: http://www.gamasutra.com/view/feature/4210/building_the_foundation_of_a_.php

[44] Wiberg, C., Jegers, K. and Desurvire, H. 2009. How Applicable is Your Evaluation Methods – Really? Analysis and Re-design of Evaluation Methods for Fun and Entertainment. Proc. of the 2009 Second International Conferences on Advances in Computer-Human Interactions, 324-328.