Identifying Onboarding Heuristics for Free-to-Play Mobile Games: A Mixed Methods Approach

Line E. Thomsen, Falko Weigert Petersen, Anders Drachen, Pejman Mirza-Babaei

Identifying Onboarding Heuristics for Free-to-Play Mobile Games: A Mixed Methods Approach

Line E. Thomsen[1], Falko W. Petersen[1], Anders Drachen[1], Pejman Mirza-Babaei[2]

Department of Communication Aalborg University, Copenhagen[1]
Faculty of Business and Information Technology, University of Ontario Institute of technology[2]
Line.e.thomsen@Outlook.dk, Falko@Weigert.dk, Drachen@hum.aau.dk,
Pejman.Mirza-Babaei@uoit.ca

Abstract

The onboarding phase of Free-to-Play mobile games, covering the first few minutes of play, typically sees a substantial retention rate amongst players. It is therefore crucial to the success of these games that the onboarding phase promotes engagement to the widest degree possible. In this paper a set of heuristics for the design of onboarding phases in mobile games is presented. The heuristics are identified by a lab-based mixed-methods experiment, utilizing lightweight psycho-physiological measures together with self-reported player responses, across three titles that cross the genres of puzzle games, base builders and arcade games, and utilize different onboarding phase design approaches. Results showcase how heuristics can be used to design engaging onboarding phases in mobile games.

Keywords: Mobile Games • Flow • Psycho-physiological Measures • User Experience • Heuristics • Onboarding Phase • Free to Play • Game User Research

1 Introduction

The Free-to-Play (F2P) business model has become increasingly popular when it comes to mobile games, and currently dominates the market. However, thousands of F2P mobile games exist, suffering from discoverability issues, and are characterized by imbalances in player retention[1]. Appsee [2] noted that the retention rate of mobile games is relatively low and only 28.60% of players return to the game after one day[2]. The reasons for the typical low retention rates varies, but one of the four most common reasons is a poorly designed onboarding experiences [2]. The onboarding phase is thus an important element for capturing and retaining new players but to the best knowledge of the authors, no heuristics have yet been developed for purpose of designing the onboarding phase. This paper will therefore focus on developing onboarding heuristics for mobile games in order to provide the industry with a set of guidelines worth considering when designing a mobile game. The exact duration of the onboarding phase of a F2P mobile game varies across titles, but in collaboration with the developers at King and Norsfell, seven minutes were established as a generic timeframe among the evaluated games. Broadly speaking, the onboarding phase lasts from the first time a player starts the game until the basic mechanics have been utilized.

The presented heuristics were formed based on an experimental study that evaluated three mobile titles: Candy Crush Jelly Saga (King, 2016), WinterForts and Pogo Chick (Norsfell, 2016). Two lightweight psycho-physiological measures, GSR and PPG were combined with abbreviated versions of the Flow State Scale (FSS) [3], using this as a proxy measure of engagement, as well as post-session walkthroughs and user-generated engagement graphs. These data were combined to develop an understanding of the features of the three games which influence flow and user experience, and jointly permitted the definition of heuristics that focus on fostering engagement of players during the onboarding phase. Because this is a short paper, the focus is on the study's results rather than the utilized method.

1) *Candy Crush Jelly Saga* is a puzzle game where the objective is to match three or more pieces of candy in a row, in order to spread jelly and hereby win the levels. **2)** *WinterForts* a base builder game where the game objective is to build, maintain, construct alliances and attack enemies. **3)** *Pogo Chick* is an arcade-style game where the player controls a chick on a pogo stick through a track of obstacles. The onboarding phases of the three titles are highly diverse as they use three onboarding styles.

Through the study of related work, the focus was on research concerning heuristics in games, psycho-physiological methods in Games User Research (GUR) and the use of Flow theory in game evaluation. **Heuristics in games:** Within the context of digital games, multiple sets of heuristics have been developed and related studies suggest that heuristics are excellent at identifying playability issues in prototypes [4]. However, during a comparison of two set of playability heuristics, Korhonen et. al [4] found that the amount of heuristics should not be too vast as redundancy then might occur and that they should be presented similarly to Nielsen's usability heuristics [4]. **Psycho-physiological measures in games:** The domain of psycho-physiological research has been described as being concerned with trying to decipher mental processes, by measuring the body's signals and thereby finding links between them [6]. Emotions are a core element in experience and using psycho-physiological measures to recognizing a user's emotional state, is utilized across fields that share related psycho-physiological states, e.g. arousal [6]. While a full breakdown of all relevant related work on psycho-physiological measures in GUR is out of scope here, a mention is made of work that utilizes the same measures in surrounding fields. In a similar study, using a Cinema setup, [8] used ECG and GSR to measure implicit psychological constructs and found physiological observations are better suited for observation of e.g. archetypes, than introspective reports. PPG is like ECG a measure that equally can be used to calculate an HRV score, which is a widely used scale to measure arousal. However, determining whether a psycho-physiological response is related to an event or non-specific can be a challenge due to the many-to-one relationship[9]. **Flow:** Flow theory was developed by Csikszentmihalyi [10]. The focus of the work, is the relationship between intrinsically motivated behavior and the associated satisfaction from performing actions in their own right, as compared to the accomplishing of a goal. According to Csikszentmihalyi the concepts of Flow consist of nine dimensions[10]. These nine dimensions are adopted in the FSS, originally developed for measuring Flow in sports [3], but adopted for use in digital games by Kivikangas and further validated by Nacke et al. [11]. The FSS measures the nine dimensions of Flow, via 36 questions. Related surveys developed specifically for

games, are targeting the user experience more broadly than Flow but still includes elements of the nine dimensions. The *Game Experience Questionnaire* [12] and *The Game Engagement Questionnaire* [13] are psychometric instruments for game user experience assessment. However, the FSS is the only frequently used questionnaire utilized for the evaluation of Flow experience [11]. A short versions of the *Game Experience Questionnaire*, termed the *Post-Game Experience Questionnaire* (PGQ) [12], covers the user experience dimensions of Positive and Negative effect, Tiredness and the feeling of Returning from a Journey. The PGQ was adopted here as a proxy measure of the user experience.

2 Method and data analysis

In brief, the participants played the onboarding phase of three different mobile games, which possess different onboarding styles. The participants played the onboarding phase of each game, without prior experience with the games. GSR and HRV data was collected during the play session, while questionnaires, interviews and engagement graphs data was collected post play. Psycho-physiological measurements measures were used as one interest was in knowing if these could be used to evaluate engagement and experience of mobile games. The experimental setup was based on a within-subject design where each participant was exposed to three different onboarding phases.

The participants included 21 females (75%) and 7 males (25%). The age of the participants ranged from 20-37 years (mean = 25.25; SD = 3.63). Demographic data was collected as well as information on prior game experiences.

Skin conduction levels were measured using a Bitalino GSR sensor. Interbeat interval was recorded by a Merlin-digital Heart Rate Monitor. An IPad mini 3 was used for the test sessions. Before each play session a three-minute baseline was recorded, where after the participants played the onboarding phase of one game. The post-play sessions consisted of multiple steps: a) drawing first Engagement Graph; b) FSS questionnaire; c) stimulated recall; d) drawing second Engagement Graph; e) PGQ questionnaire. This approach was followed for each onboarding phase condition.

Data analysis consisted of multiple steps: a) post-session processing of the psycho-physiological data; b) evaluation of the FSS and PGQ surveys; c) explorative content coding of the post-session interviews; d) evaluation and comparison of the engagement graphs with GSR and HRV composite graphs; e) comparative and correlational analysis across the qualitative and quantitative measures. The GSR data was first visually inspected for logging gabs, where after the data was normalized using an algorithm that was used in similar analysis by Mirza-Babaei et al. [14]:

$$GSR_{Normalized} = \frac{(GSR_{(t)} - GSR_{Min})}{(GSR_{Max} - GSR_{Min})} \qquad (1)$$

GSRMin and GSRMax refers to the minimum and maximum GSR value in a certain time frame and *GSRt* is the GSR data point in the middle of the time frame [14]. The GSR and HRV data was divided into smaller time windows that were based on the

game design. Based on these time frames, graphs were created to visualize the normalized GSR and HRV scores. The GSR and HRV graphs were compared to the engagement graphs, created by the game designers and participants.
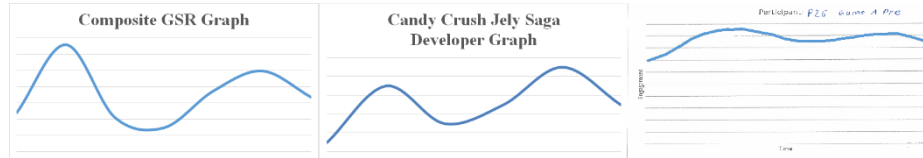
Fig. 1. Example of the three different graph types, Composite GSR(Left), Developer Engagement Graph (Middle), Player drawn Engagement Graph (Right).

In order to check the reliability of the PGQ and FSS, Cronbach's alpha was calculated. For the FSS the average score was .64 (SD = .42) while the PGQ had an average score of .71 (SD = .13). In order to investigate which of the three games induced the highest Flow score, the average score across all Flow dimensions was calculated. A comparison of overall Flow and PGQ scores is displayed in Table 1.

Table 1. Average FSS score for the three tested games and the PGQ dimension positive and negative game experience

Game	Average Flow scores	PGQ Positive Dimension	PGQ Negative Dimension
Candy Crush Jelly Saga	3.84 (SD=1.03)	2.27 (SD=1.19)	1.52 (SD=.94)
WinterForts	3.11 (SD=1.26)	1.74 (SD=.95)	2.04 (SD=1.40)
Pogo Chick	3.04 (SD=1.30)	2.08 (SD=1.40)	1.83 (SD=.99)

The video recordings were meaning condensed and then coded by the means of an open coding. The engagement graphs were further analyzed by comparing them to the statements given during the interviews. The sentiments describing the participant's experience of the onboarding phase and thus the graphs, were used to analyze which game element created an enjoyable onboarding phase. Based on the analyzed data several onboarding heuristics were identified.

3 Identifying Onboarding Heuristics

The following top level categories were identified through the open coding: 1) Game context, 2) Music, 3) Learnability, 4) Rewards System 5) Relatedness 6) Mastering Controls 7) Game Objective and 8) Progress Overview. The categories 1,2 and 3 seemed to be related to the device being mobile and the context thus ever-changing. The players were e.g. conscious of not disturbing their surrounding while playing and also needed games that can be played despite of interruptions caused by their surroundings. 4,5 and 8 are related to the games providing valuable and meaningful feedback for the players in order for the players to be able to relate to a game they never played before. Lastly, 6 and 7 are related to the game mechanics, where the controls need to be manageable and the games objective needs to be clear to the player. Additionally,

several of the categories identified (4, 6, 7 and 8) in the open coding showed a shared relationship with the nine dimensions of Flow and a selective coding was therefore performed to investigate if more categories would uncover. The selective coding uncovered the supplementary categories Immersion and Autonomy. The categories 6, 7 and 8, initially found during the open coding, were refined and renamed: Skill Level, Clear goal and Clear feedback and Progression.

The heuristics identified in the empirical data will know be presented. The description of each heuristic follows the style of Nielsen's [4]. The heuristics are: **Clear goal**: Players should have a clear understanding of the games goal; this can be established in the beginning or during gameplay. **Autonomy:** Players should feel in control when playing a game. Missing autonomy can cause a negative game experience. **Skill Level:** The game should match a player's skill, without being too hard or easy. **Relatedness:** Players need to feel an association with the game. The player thus needs to feel that e.g. the theme, visual design and genre of the game is relatable. **Clear feedback and progression:** Players should receive information about how well they are during and when succeeding in a goal. **Game context:** The game needs to allow varying contexts and for disturbances without the players making irreversible mistakes that permanently ruins the player's game experience. **Learnability:** Players should be able to learn core game mechanics quickly during the onboarding phase, to ensure that they know how to play the game. **Music:** Music in the game can help support the game genre and environment but usages of sound depends on the surrounding context, as the players can be cautious about causing disturbance. **Rewards system:** Rewards needs to have value and purpose for the player. Through the onboarding phase the player needs to learn the reward system and find the rewards desirable to obtain when playing. **Immersion:** Immersion can be the outcome of a well-designed onboarding phase.

4   Discussion and Future work

The goals of this study was to identify which elements that create an enjoyable onboarding phase, and based on this identify heuristics. Furthermore, lightweight psycho-physiological measures were utilized to investigate if it was possible to utilize the techniques in a mobile game setting. It was concluded that these measures of arousal appear to correlate with the self-reported data and that the high-frequency measures can be used to obtain insights into the specific causes of arousal. This result parallels previous work in other types of games [11, 6]. During the evaluation of the three games, the most common problem identified was a lack of autonomy. When players were forced to perform certain actions, the participants expressed a negative game experience. Furthermore, some of the identified heuristics are share with related literature (e.g. Clear Goal and Autonomy) on playability heuristics [4],which indicates that the heuristics, presented here, might be more generic across games and platform. Even though several sets of heuristics have been developed, none of them are focusing on the onboarding phase of mobile games. The heuristics thus contributes with knowledge within a relatively undiscovered research area. As the presented heuristics still are at an early stage, future work will focus on testing and evaluating them in order to create more refined

and effective heuristics, which enables the development of better onboarding phase and increases player retention rate. Additionally, this study indicated that lightweight psycho-physiological and self-reported data can support each other in the evaluation of the onboarding phase of mobiles games, as they can pinpoint even small game events that causes high or low arousal, which will be elaborated further in the future. In conclusion, the three tested games had very different onboarding phases, which helped identifying the elements that created a more or less enjoyable experience for the players. This formed the basis for identifying the 10 onboarding design heuristics. However, future research has to be conducted to refined and evaluate the heuristics further.

References

1. Hadiji, F., Sifa, S., Drachen, A., & Thurau, C. Predicting Player Churn in the Wild. Player Churn in the Wild. In Proceedings of the IEEE Computational Intelligence in Games. pp. 1-8 (2014).
2. Appsee Insights. Retrieved at:http://blog.appsee.com/blog/2015/10/13/gaming-app-user-retention-only-22-return-after-one-month/
3. Jackson, S. A., & Marsh, H. W. Development and Validation of a Scale of Measure Optimal Experience: The Flow State Scale. Journal of sport & exercise psychology. 18, pp. 17-35 (1996)
4. Korhonen, H., Paavilainen, J., & Saarenpää, H. Expert review method in game evaluations: comparison of two playability heuristic sets. Proceedings of the 13th International MindTrek Conference, pp. 74-81 (2009)
5. Mirza-Babaei, P. Biometric Storyboards: A Games User Research Approach for Improving Qualitative Evaluations of Player Experience. Sussex: University of Sussex. Doctoral thesis (PhD), University of Sussex (2013)
6. Ivonin L., Chang HM., Díaz M., Català A., Chen W., Rauterberg M. (2015). Traces of unconscious mental processes in introspective reports and physiological responses. PLOS ONE, vol. 10, no. 4, e0124519, pp. 1-3 (2015)
7. Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G.The Handbook of Psychophysiology. New York: Cambridge University Press. (2007)
8. Csikszentmihalyi, M. Beyond boredom and anxiety. Jossey-Bass: San Francisco. (1975)
9. Nacke, L., & Lindley, C. A. Flow and Immersion in First-Person Shooters: Measuring the player's gameplay experience. In proceedings of the 2008 Conference on Future Play: Research, Play, Share. pp 81-88 ACM: Toronto (2008)
10. Isselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, J. Characterizing and Measuring User Experiences. In proceedings of ACE Conference '07. In proceedings of The International Conference on Advances in Computer Entertainment, vol 3(8) pp 33-39 ACM: Toronto (2007)
11. Brockmeyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. The development of the Game Engagement Questionnaire. Journal of Experimental Social Psychology, vol. 45 pp. 624-635. Oxford University Press, Oxford (2009)
12. Mirza-Babaei, P., Wallner, G., McAllister, G., & Nacke, L. E. Unified visualization of quantitative and qualitative playtesting data. CHI '14 Extended Abstracts on Human Factors in Computing Systems pp. 1363-1368. ACM. New York (2014)