



Understanding students' game experiences throughout the developmental process of the number navigation game

Phuong Bui¹ · Gabriela Rodríguez-Aflecht¹ · Boglárka Brezovszky¹ · Minna M. Hannula-Sormunen¹ · Samuli Laato^{1,2} · Erno Lehtinen¹

© The Author(s) 2020

Abstract

Serious games for learning have received increased attention in recent years. However, empirical studies on students' gaming experiences throughout the developmental process of serious games and discussions regarding game design are missing. The aims of the present study were to analyze students' gaming experiences while playing four consecutive versions of the Number Navigation Game (NNG)—a mathematical game-based learning environment focusing on flexibility and adaptivity with whole-number arithmetic; and to provide an extensive review of the NNG developmental and design process over 3 years with focus on how and why the design decisions were made, and how those choices affected students' gaming experiences. The study employed a mixed-methods design of quantitative and qualitative research. The Game Experience Questionnaire about eight core game experience dimensions was answered by different groups of students at primary schools in Finland in three different experiments after students played four versions of the NNG from 2014 to 2016. Six semi structured interviews related to students' game experiences, preferences and game features of the latest version of NNG were conducted. Overall, results indicate that improvement in game's usability and clarity in the user interface has positive impacts on students' game experiences. Furthermore, there seems to be a clear advantage in having better aesthetics and value in improving extrinsic elements that could contribute to maintain players' enthusiasm and situational interest in serious games.

Keywords Game-based learning · Serious games · Game experience · Serious game design

✉ Phuong Bui
phuong.bui@utu.fi

¹ Department of Teacher Education, Center for Research on Learning and Instruction, University of Turku, 20014 Turku, Finland

² Department of Future Technologies, University of Turku, 20014 Turku, Finland

Introduction

Over the past decades, digital games have emerged as instructional tools for educational purposes. Game-based learning has received exceptional attention owing to the evidence of positive impacts on learning outcomes at all levels of education (Ma et al. 2012; Wouters et al. 2013) and their potential to engage students through meaningful and challenging tasks (Whitton 2011). Despite much effort, initially educational games often fell into the piles of “chocolate-dipped broccoli” as famously termed by Bruckman (1999) because they were neither engaging nor fun. Eck (2006) argued that game-based learning design needs to find “synergy” between educational goals and engaging factors; whereas Habgood and Ainsworth (2011a; b) advocated for designing intrinsically integrated games, in which learning materials are embodied within game mechanics.

Designing and developing such games is a complicated and challenging task. Game design must balance learning and fun, appeal to as many players as possible without compromising the educational benefits (Kiili et al. 2014). In other words, an effective serious game needs to have a sound pedagogical framework and intriguing gaming elements. Such educational games, however, were often inadequately interpreted as the focus mostly fell on whether the games fulfilled their educational promises, rather than how they were designed and developed, how the design choices were made (and why), and how those choices (including mistakes) affected the outcomes (Gaydos 2015). This is in alignment with works on designing and integrating purposeful learning in gameplay by Ke (2016), which pointed out that previous studies on game-based learning have mainly focused on reporting the effectiveness of games and did not provide detailed descriptions of game design features and processes. Indeed, concerns about the quality of serious game design are also often absent in serious game studies (Mitgutsch and Alvarado 2012).

The more engaging the gameplay experience, the more likely learners want to play the game (Birk and Mandryk 2013; Oksanen 2013). Furthermore, understanding players’ experience is a key factor that help in figuring out whether players are engaged in and motivated to play the games (Hamari and Kehonen 2017). However, there is a lack of empirical studies about students’ gaming experiences in authentic settings throughout the developmental process of game-based learning environments (De Grove et al. 2010). Moreover, not all game studies provide comprehensive descriptions of games’ major characteristics with discussion of changes from prototype to software trials to final forms (Habgood 2007; Oksanen 2013); which makes analyzing and understanding the gaming features and activities during gameplay more difficult (Torbeyns et al. 2015). This again reinforces Gaydos (2015) and Ke (2016)’s arguments that we should shift the emphasis on the developmental processes of serious games, in which theoretical underpinnings, design strategies, rationales for game features are explained and documented. From a game design perspective, game designers need to pay attention to how even small changes of the game can affect others (Hunicke et al. 2004), and how those changes might have different impacts on the gaming experience and learning process.

This study analyzes students’ gaming experiences playing four different versions of the Number Navigation Game (NNG) in school settings over the course of 3 years to understand how students’ gaming experiences varied between versions of the game. Previous studies with the NNG show that students’ mathematical skills improved after they played NNG (Brezovszky et al. 2015), and students who practiced more with NNG also benefited more from the gameplay, meaning that extended practice with NNG was able to develop students’ recognition and use of numerical characteristics and relations (Brezovszky et al.

2019); however, motivation for math learning slightly decreased, and gaming experiences were negative (Rodríguez-Aflecht et al. 2015). These results suggested that changes that improve gaming experience and strengthen motivation were needed.

Why the number navigation game?

NNG is a mathematical game-based learning environment developed to improve the mathematical skills of primary school students, specifically focusing on flexibility and adaptivity with whole-number arithmetic (Brezovszky et al. 2015; Lehtinen et al. 2015; McMullen et al. 2017), and at the same time triggering and maintaining students' interest in mathematics learning (Rodríguez-Aflecht et al. 2018). NNG has an "intrinsically integrated" design (Habgood and Ainsworth 2011a, b) in which the core gaming mechanism is integrated directly to the educational content of the game. For detailed discussion of NNG development and roles of different game features in enhancing adaptive number knowledge, see Brezovszky (2019).

The development of NNG was a collaborative process carried out by a multidisciplinary team composed of researchers, educators and programmers with varied but complimentary experiences and qualifications covering different learning, design and technical aspects (mathematics, motivation, measures of learning outcomes, game development, etc.). NNG went through continuous circles of testing and development: the earliest prototype and pilot from a usability perspective (Brezovszky et al. 2013) to a large-scale, randomized, controlled study using a more advanced version with few motivating elements (Brezovszky et al. 2019; Rodríguez-Aflecht et al. 2015). Through this process, NNG expanded from two-dimensional graphic versions with limited focus on design aspects to a three-dimensional graphic version where game components were purposefully designed and more extrinsic motivating features were available.

Games with intrinsic integration are argued to be more effective in achieving learning outcomes and higher motivation than their extrinsic equivalents, however, they can be more expensive and difficult to develop (Habgood and Ainsworth 2011a, b). Balancing the learning content and game design would demand continuous testing and development activities, which requires extensive research fund and usually hinders the application of research-based games into the real world when funding ended (Gaydos 2015). Findings in the field of game-based learning specifically suggest that there is a strong need for more research with longitudinal study designs (Hainey et al. 2016; Young et al. 2012) and focus on game design, instructional design and game features' effects in learning (Wouters and van Oostendorp 2013).

Hence, describing and documenting the developmental process of an intrinsically integrated game like NNG with focus on how and why the design decisions were made, and how those choices affected gaming experiences would contribute greatly to the understanding of educational game design. The development of NNG could serve as an example of an iterative and multidisciplinary design and development process, where design choices were based on theory and empirical results (Vanden Abeele et al. 2012). Furthermore, it also provides the opportunity for others to understand which features were needed to create a meaningful serious game, and how some changes might or might not be as impactful as intended. Finally, the earlier versions of NNG were included very few extrinsic elements with simple design and aesthetics, whereas the final product NNG 4 has similar intrinsic integration but also many more extrinsic elements. This provides a rare chance to

investigate whether there is a clear advantage in having more extrinsic elements and better aesthetics in game-based learning or not.

Understanding the serious game experience

Many studies have applied flow theory (Csikszentmihalyi 1975), regarded as the central concept in user experience, as a framework for modeling enjoyment, engagement, and satisfaction. The idea is that when one enters the “flow state,” one experiences an “optimal experience” where one is so engaged with the given activity that nothing else seems to matter. Studies have pointed out that being in the flow state has positive effects on learning (Engeser and Rheinberg 2008; Kiili and Lainema 2008). Findings also have shown that flow influences players’ enjoyment and performance (Weibel and Wissmath 2011).

Immersion is another concept that is linked to engagement, and broadly considered as outcome of a good gaming experience. According to Ermi and Mäyrä (2005), immersion can be divided into three components: sensory, challenge-based, and imaginative. Brown and Cairns (2004) suggested that to be engaged with the game, players needed to invest time and attention in learning how to play and control the game. Then, if players were further involved with the game, they might feel “engrossed” and became less aware of their surroundings or themselves. The last level of immersion is “total immersion,” the highest level of attention, when players are so immersed to the game that they “stop thinking about the fact that (they) are playing a computer game” (Brown and Cairns 2004). Flow theory and immersion are similar as they indicate a sense of losing oneself in the game or focusing solely on the task at hand without paying attention to the time or one’s surroundings. Jennett et al. (2008) argued that flow is an extreme sort of experience; whereas immersion is an experience that progresses through varying degrees of engagement, and one can be immersed into the game without entering the “flow state.” Thus, flow might be the optimal end, or the “total immersion” level. Kiili et al. (2012) also considered immersion a “lower level” expression of flow, explaining that flow happens when a player directs all of his or her attention to a given task, and immersion is when a player “physically or virtually becomes a part of the experience itself” (Kiili et al. 2012, p. 85). Others, however, defined the two concepts somewhat differently, as they do not believe “flow” and “immersion” are two degrees of the same thing. Oksanen (2013) argued that immersion refers to one’s sense of presence in a mediated environment, and flow is the involvement in an activity. This is in alignment with Weibel and Wissmath’s (2011) findings, in which the experience of flow referred to the perception of being highly involved in the gaming action. These concepts, nonetheless, are not broad enough to encompass the multidimensional and complex concept of gaming experience. Other factors also contribute to gaming experience, and there are different ways to model the core elements of the gaming experience. Poels et al. (2007) proposed a framework consisting of seven dimensions of game experience as (a) flow, (b) (sensory and imaginative) immersion, (c) competence, (d) challenge, (e) positive affect, (f) negative affect, and (g) tension (cf. Burnes et al. 2015). Thus, within this study, Poels et al.’s (2007) framework is used to explore game experience, as this framework is more suitable for the nature of this research. Competence and challenge are highly connected to flow and immersion. Csikszentmihalyi (1990) explained that the balance between challenge and ability is one of the key components of the flow state. Others have confirmed that a match between the player’s skills and the game’s challenges is a precondition for flow to happen (Kiili and Lainema 2008; Sweetser and Wyeth 2005). Other dimensions

are positive affect, negative affect, and tension. They are indicators of how enjoyable game experience is for the player. Positive affect means general positive emotion, such as fun and enjoyment; negative affect refers to boredom or lack of concentration. Last, tension is deeply related to negative affect; however, tension concerns stronger emotions, such as irritation or frustration.

The last dimension discussed in this paper is positive value. The idea is that players need to first believe that playing a serious game would be helpful to them before they can benefit from the game-based learnings (Whitton 2011). A previous study of the NNG considered that a belief in the positive value of the game complements Poels et al.'s (2007) framework, which measures students' belief that the NNG is helpful to them (Rodríguez-Aflecht et al. 2015). Therefore, this study would not be fully complete if positive value dimension were not measured.

Research tasks and aims of the study

This paper describes the development and testing of four game versions (NNG 1-NNG 4) across three studies. The overall aim of these studies was to present a comprehensive description of the NNG developmental process and to understand the gaming experiences of students playing different versions of the NNG. As between NNG 3 and NNG 4 there were major changes to the extrinsic elements of the game, the second goal was to explore in detail students' preferences and attitudes towards the improvement of the NNG's extrinsic elements and overall visual aesthetics. General aims of this paper are to answer the following research questions.

Research question 1

How do changes in the different versions of NNG affect players' gaming experiences?

Rodríguez-Aflecht et al. (2015) conducted a study on students' gaming experience playing NNG version 1 (NNG 1). Based on this study, our hypothesis is that the game mechanism of NNG is beneficial for learning, but other elements are not motivating or engaging enough. Thus, different changes were made (see Table 1) in NNG 2 and then NNG 3. This study examined students' experiences playing three different versions of NNGs and investigated whether improvements in external motivating elements and user interfaces result in different experiences among participants; and if there are different experiences, how different they are and in what ways.

Research question 2

What are the students' attitudes and preferences towards the changes and features in NNG 4 compared to NNG 3? How are changes in NNG 4 related to students' attitudes and intention to play the game?

In NNG 4, major changes were made on different extrinsic aspects of the game: game structure, motivational elements, and aesthetic improvement. Aesthetics are often considered essential elements in video game experiences, and it is argued that good aesthetics can compensate for imperfections in game design and attract players to games that they might have ignored (Schell 2014). Thus, we expected to understand more about students' attitudes and preferences towards those changes in NNG 4 and how the

Table 1 Summary of changes and updates during the developmental process of NNG

	NNG 1	NNG 2	NNG 3	NNG 4
Game mechanics	Two game modes (scoring and energy)	Similar game modes Introduction of the pirate ship and hidden operations	Similar to NNG 2	Similar game modes Pirate ships appear more frequently compared to previous versions
In-game user interface	Displayed minimal information about the location of the ship and the material Players used both keyboard and mouse to enter to operations Game situation was saved after the whole map was completed If students closed the game within a map, they had to repeat all the already played materials of the map in the new game session	Added a number pad Introduced visual cues (green mode, blue and footprints for the scoring mode) and the materials bar Game situation was saved after each material was in the harbor. When students reopened the game, they did not have to repeat any previously played materials but were able to continue from the point they were when closing the previous game session	Similar to the NNG 2 design	Sidebar was redesigned to be more visually attractive and related to the ship exploration theme Visual feedback with clear cues Clear task instruction and always visible on top of the number pad
General game user interface	Level overview window had minimal information about the game status: number of coins collected, latest map completed Map window increased the number of maps (4 to 12, 12 to 20, 20 to 28) each time after players complete the all the available maps in the previous window	No significant changes were made	Introduced more coherent game levels: increased the number of levels (8 levels) but decreased and standardized the number of maps per level (6 maps per level) Added information about the currently used ship, currently played map, game mode, and total earned coins	The map window is islands expanding on the design of a treasure map. Instead of numerical order levels, players explore island to island Information about players' earned coins, accumulated materials, ships, etc., can be accessed either via the Treasure Box icon or on Lighthouse Island

Table 1 (continued)

	NNG 1	NNG 2	NNG 3	NNG 4
Help functions	No built-in dynamic help function	No built-in dynamic help function	Introduced a Help page that contained general instructions for the game rules and strategies	In-game tutorial provides step-by-step instructions for how to play the game, including introduction of the game's interface, rules, and elements Introduced the Hint option, which highlights numbers where the boat can be moved to after any possible operation
Motivational mechanisms	Players earned coins (bronze, silver, or gold) when they finished a map The idea is players need to collect all 4 materials to build the village (which is displayed as an icon on the sidebar) The materials earned during a map stayed on the map	Players earned coins (similar to NNG 1) Instead of a village icon, upon retrieving the materials, a pop-up window displayed a large picture of the village changing from basic to more modern The materials earned during a map stayed on the map	Players earned coins (gold) to build the village similar to NNG 2 A Shop was added where players can buy and upgrade ships using the coins they earned The ships were shown from basic to more modern ones. All ships functioned the same way The materials earned during a map stayed on the map	Players earn gold coins similar to NNG 3 The village window and Shop were abolished. Instead, a Lighthouse Island was integrated, where players can construct buildings and buy ships using materials and coins they earned from completing maps
Customization	Players had no customization option	Players had no customization option	Players had no customization option	Players can customize some aspects of the user interfaces: language setting, game setting (displays of the number pad, sound effects, music), avatar customization, personal name, saving, loading and replay gameplays

Table 1 (continued)

	NNG 1	NNG 2	NNG 3	NNG 4
Visual appearance	2-dimensional graphics with simple design Did not have a clear theme or storyline	Some visual aspects were improved compared to NNG 1 (coin display, material icon, sidebar, etc.)	Similar design to NNG 2 Shop displayed different kinds of ships (from basic to modern); the theme ship exploration was clearer	3-dimensional graphics with colorful, vivid animations. The game is more visually attractive and similar to commercial games The theme ship exploration is strongly apparent in all components of the game

improvement related to students' attitudes toward the game and their intention to play the game through a group of students who had played both NNG 3 and NNG 4.

The number navigation game

The core game mechanisms described here used an NNG 1 map as an example, but they remained consistent throughout the developmental process. The game interface is 100 squares superimposed on various maps of land and sea, where players are given tasks of collecting four different raw materials to build settlements. Players progress by navigating a ship from a starting point (the harbor) to retrieve a material on a given point and return to the harbor by applying various combinations of numbers and arithmetic operations. For instance, in Fig. 1, the player starts from number 89 and has to collect wood situated at number 62. The player has to move by inputting mathematical equations on the left side of the screen. The moves have to take the ship to the targeted material (number 62) and avoid numbers covered by land (Brezovszky et al. 2015). Players completed a map by collecting all four materials.

Generally, NNG has two scoring modes: moves and energy scoring. The principle of the moves scoring mode is simple: players need to retrieve the materials and return to the harbor using the least number of moves possible. For the energy scoring mode, the idea is to use the minimum energy, because the energy is measured by adding up all the numbers inputted in the operation box.

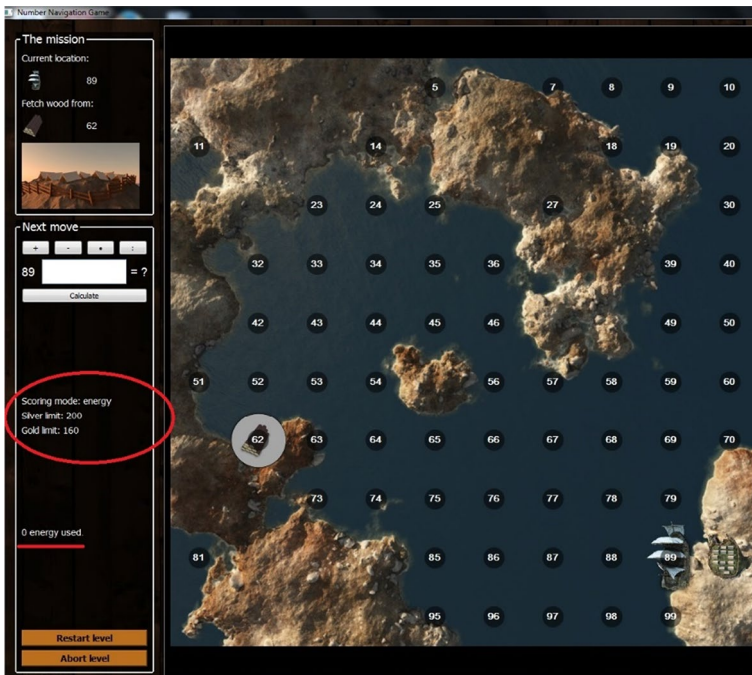


Fig. 1 Example of a map in the energy scoring mode in NNG 1

NNG 1 and NNG 2

Differences between NNG 1 and NNG 2 were the usability and clarity of the game interface. A Number Pad was added allowing players to use either keyboard or mouse. As in NNG 1, it would be quite frustrating when players had to shift from using mouse to keyboard to enter the equations. Visual cues and materials bar were introduced to inform players about the status of the map. In NNG 1, game situation was saved only after the map was completed, which could be irritating as players had to finish the whole map to save their progress. Therefore, in NNG 2, game situation was saved after each material was in the harbor. Additionally, the pirate ship and hidden operations were added in NNG 2 to create more variability in the use of numbers and operations. When the pirate ship appears, players cannot reserve the operations that took them to the material. Hidden operations refer to the operations that are not available on some maps; this encourages players to look for less obvious solutions for a problem. These features were motivated by the premise that to develop adaptive number knowledge you need a large amount of practices with various different numerical relations and number combinations (see Brezovszky et al. 2019, 2015; McMullen et al. 2017); and based on results of previous studies which showed that the mechanical repetition of inverse operations could be problematic after a while (see Lehtinen et al. 2015; Rodríguez-Aflecht et al. 2015).

NNG 3

Based on empirical findings, observational data, and feedback from users (Lehtinen et al. 2015; Rodríguez-Aflecht et al. 2015), NNG 3 was developed in 2015 (Fig. 2). An arrangement of difficulty levels and new motivating mechanisms were introduced.

Reward system A Shop was added, where players could buy and upgrade ships using coins earned during gameplay. While the game was efficient in developing the desired math learning outcomes, its mechanism were not engaging enough to keep up math motivation in all students (Rodríguez-Aflecht et al. 2015). Thus, adding the Shop was expected to give meanings to the coins that players earned via completing maps (Lehtinen et al. 2015), and give incentives for players to complete more maps to earn more coins.

Order of difficulty levels and maps Previously, total number of maps was 64, and the number of maps per level increased from 4 to 12, 12 to 20, and 20 to 28 maps (Lehtinen et al. 2015). Organizing levels like that could be overwhelming as players might have stuck only to easier maps or moved maps often when there were many available. Therefore, changes for better control over level were added. There were eight difficulty levels with six maps per level (Fig. 3), and the number of maps per level remained consistent, totaling 48 maps.

The sidebar displayed immediate feedback mechanisms and inform players about the status of their gameplay. Once a map was completed, earned coins appeared over the map's thumbnail; there were no coins over uncompleted or new maps.

Help functions Help page contained texts and screenshots regarding game's rules and strategies. This was needed as understanding the rules was sometimes problematic, and players could get help even when/if teachers were not available.



Fig. 2 Example of a map in NNG 3

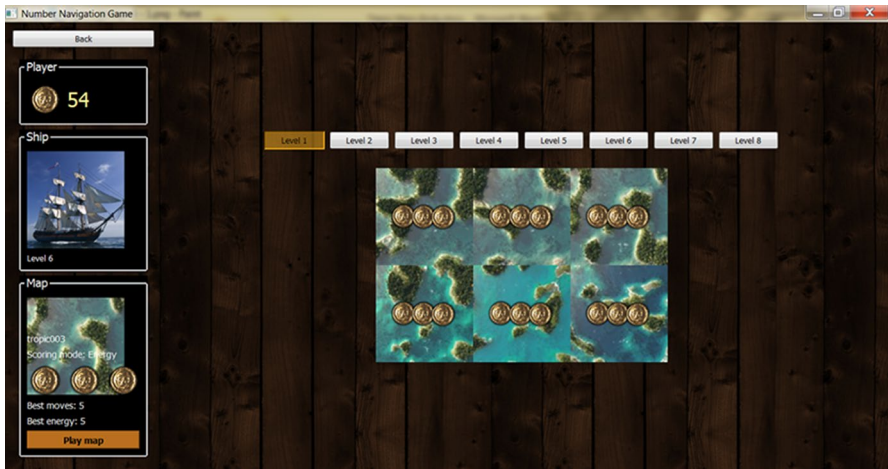


Fig. 3 Arrangement of maps in NNG 3

NNG 4

The development of NNG was continued based on results from previous studies (Brezovszky et al. 2015; Lehtinen et al. 2015; Rodríguez-Aflecht et al. 2015). NNG 4 was developed with more extrinsic elements, and the testing version was released in spring 2016 (see Table 1).

Visual appearance NNG 4 was developed with the Unity3D game engine instead of QT. A clear theme—Ship Exploration was visible in all design components. These design elements were carefully chosen and holistically relate to one another. Graphic improvement



Fig. 4 Main screen of NNG 4



Fig. 5 Lighthouse Island and where players can purchase ships

is expected to make the game more visually attractive and contribute in creating a fantasy world that previous versions lacked (Fig. 4).

Motivational mechanisms A Light House Island was introduced (Fig. 5). In previous versions, materials earned during a map stayed on the map, and it was implied that those materials were used for building the villages (as a pop-up window displayed an image of a village changing from basic to more modern). In NNG 4, materials were accumulated throughout maps, and players use them to construct buildings on the Light House Island. Similar to NNG 3, players use coins to purchase ships. However, in NNG4, a new feature was introduced which allowed players to pick up multiple resources between a starting and a target number. This feature was important as it required players to perform more complex mathematical problem solving, and it also gave meaning to ship upgrades as only bigger ships could carry multiple resources at the same time.

Order of difficulty levels and maps In NNG 4, the map window (Fig. 6) was inspired by an archipelago island layout expanding on the design of treasure maps. Instead of playing Level 1, Level 2, etc., players advance in the game by unlocking islands.

Help functions In NNG 4, players started playing with a tutorial that give step-by-step instructions for the first gameplay (Fig. 7). Students and teachers involved with previous versions learned about game rules and aims via a separate online video (Lehtinen et al. 2015), usually with support of trained researchers (Brezovszky et al. 2019; Jaatinen 2016). In NNG 3, a separate Help page was added. However, that was no tutorial. Players can play the tutorial unlimited times and they can go back any time. Another new feature was the Hint option



Fig. 6 Map window in NNG 4



Fig. 7 The parrot in the tutorial mode

(Fig. 7) indicated by highlighted numbers on the map. These are possible spots that the ship can be moved to after any operation.

Customization In NNG 4, players can customize some user interfaces: language setting, avatar customization, saving, loading, and replaying games. These customizations do not affect the game mechanics and dynamics directly.

Methods

Research design

This study reflects the natural process of developing a research-based educational game, where it makes use of data collected from different studies within a larger project, each with different designs depending on the research questions of each study. This study applied mixed-methods design -a procedure that takes advantages of quantitative and qualitative methods (Tashakkori and Teddlie 2010). Quantitative data is collected via questionnaires from previously done studies in 2014 and 2015 (see Jaatinen 2016; Lehtinen et al. 2015; Gabriela Rodríguez-Aflecht et al. 2015). The qualitative data were collected via semi-structured interviews after students played NNG 4 in 2016. Table 2 shows the information about the participants and these studies.

Table 2 Overview of studies, participants' information and NNG versions

2014	2014	2015	2016
NNG 1 4th to 6th graders N=642	NNG 2 4th to 6th graders N=526		
NNG1 4th graders' subsample N=63	NNG2 4th graders' subsample N=72	NNG3 New 4th graders sample N=40	NNG4 Same group as in 2015 now 5th graders N=37

Participants

The samples used in the study consisted of three cross-sectional data sets collected in different time points during the development phases of the NNG game. First gaming experience study was made as a part of a large-scale experimental study with aimed at exploring if the unique game mechanics of the NNG result in desired learning effects in mathematics. The sample used in the large-scale study consisted of 1168 students from 61 fourth- to sixth-grade public school classrooms across four cities and towns in South Finland in spring 2014. NNG 1 was played by the experimental group which consisted of 642 ($n=299$ girls) students. After the posttest of the experiment, the 526 ($n=247$ girls) control group students played the NNG 2 version.

Because the main focus group for which the game was developed are 4th graders, the comparison of gaming experiences between the first two versions (NNG 1 and NNG 2) and the later version (NNG 3) was made among the 4th graders' subgroups of earlier studies and the later selected 4th graders' sample. NNG 3 was played by 40 fourth graders ($n=13$ girls) from two classrooms at a public school in a city in Southwest Finland in 2015. The socio-economic and ethnic background of students was quite similar compared to samples that played NNG 1 and NNG 2.

The same students who played NNG 3 in 2015 played the prototype of NNG 4 when they were in fifth grade in spring 2016. Three of the 40 students were not at school during the NNG4 data collection.

Participation was voluntary, and informed consent forms were gathered in writing from the participants' parents. The ethical guidelines of the University of Turku were followed strictly.

Instruments

Questionnaire

Game Experience Questionnaire (GEQ) aims at measuring eight core dimensions of game experience: challenge, competence, flow, immersion, negative affect, positive affect, tension, and positive values. The questionnaire was in Finnish with translation by Oksanen (2013); it was also simplified in language and in length (see Appendix) to be more suitable for the age of participants. Among the items, 15 of the 42 items from the original GEQ in Oksanen (2013) were removed, 3 additional items of Positive Values dimension and 1 item

Table 3 Cronbach alpha reliability values of dimensions of game experience in the three NNG studies

Dimensions	Item number	α NNG 1	α NNG 2	α NNG 3
Positive values	3	.81	.87	.83
Competence	4	.81	.90	.89
Immersion	3	.78	.81	.77
Challenge	4	.65	.58	.38
Negative affects	5	.79	.85	.78
Tension	2	.77	.74	.38
Flow	5	.79	.84	.63
Positive affects	5	.94	.94	.95

of Challenge dimension were added. Each item was a statement related to game experience. A 1–5 scale indicated level of agreement, with answers from 1 (not at all) to 5 (extremely). According to Rodríguez-Aflecht et al. (2015), the factor structure of the 31 items in GEQ was studied through principal component analysis with varimax rotation; data was adequate for factor analysis with a 0.95 Kaiser–Meyer–Olkin Measure and Barlett’s test of sphericity show a significance of $p < 0.001$. Eight separate factors were found and used as basis for subscales. Reliability of the subscales by phases is presented in Table 3. Most of the reliability is sufficient, except for Challenge ($\alpha = 0.38$) and Tension ($\alpha = 0.38$) in NNG 3. Due to low reliability, these two dimensions in NNG 3 were removed in further analysis.

Semi-structured interviews

Interviews were conducted individually in English with students from a specialized English class, where students study half of the time in English and half of the time in Finnish. Six students volunteered (three boys and three girls) to participate. Students were considered to be able to discuss their game experiences comfortably without any language barrier. The design of the interview questions was based partially on the previous results about NNG (Rodríguez-Aflecht et al. 2015; Brezovszky et al. 2015; Lehtinen et al. 2015) and partially based on the development of NNG 4. Since the participants had taken part in the experiment of NNG 3, they were introduced to the NNG before and already formed some impressions about the previous version of the game. The interview sought to understand how students react to the improvement in game structure, motivational elements, and aesthetic, and how these alterations affect their attitudes and intentions of playing NNG 4. Interview questions were divided into three parts: (1) introductory scripts where students were presented with information about the interview and its purposes, (2) warm-up questions where questions related to students’ previous gaming experiences’ with NNG 3 were presented, such as “Can you tell me something about your experience playing the game?” and “What did you like/dislike about that game?”, (3) the last part is substantive questions—where students were asked questions related to their experiences with NNG 4, such as “Do you see any changes of the game? What are they?”, “Do you enjoy playing this game (NNG 4)? /What makes it enjoying/less enjoying to you?”. Follow-up questions related to students’ practical game experience were also used.

Data collection procedures

Quantitative data Participants played NNG 1 for a 10-week period as part of their regular math classes, while NNG 2 group studied with regular mathematics curriculum. Afterward, the conditions were reserved, and NNG 2 participants play for at least 10 h in total for 5 weeks, during which one session lasted at least 30 min. For more details see Rodríguez-Aflecht et al. (2015).

In NNG 3 experiment, participants played for about 10 h in total. It was suggested that students play the games three times a week in 45 min slots during math class. Before the intervention, teachers received guidance about how to use the game in their classroom practice. Students played the game individually as part of their math class activity for a 5-week period. One of the classes played the game on personal hybrid tablet PCs at the end of their math class. The other class played the NNG in the IT room with the school PCs for 30 to 40 min per session. For more details see Jaatinen (2016).

In NNG 4 experiment, students played the game in two math lessons for two consecutive weeks. Each section lasted 90 min (2 lessons/ 2×45 min). Students played the game either on their designated laptop or in the computer room.

Qualitative data The semi-structured interviews were conducted right after students played NNG 4 the second time. On average, each interview lasted about 7 min, except for one that lasted for 15 min. The interviews were planned to last from 6 to 10 min. Five participants agreed for the interviews to be taped, and one interview was taken notes. All data were later transcribed for analysis.

Data analysis

Quantitative data

Data gathered from three experiments using identical post play GEQ including 31 variables accounted for eight core dimensions of game experience. After the reliability test (Cronbach's alpha), sum variables were calculated for the subscales. Game experiences with participants from 3 grade levels (4th to 6th grade) between NNG 1 and NNG 2 were tested with an independent-samples *t* test; while a one-way ANOVA was carried out to compare 4th graders game experiences playing three different versions of NNG from NNG 1 to NNG 3.

Qualitative data

The method of analysis elected for the qualitative data was Braun and Clarke's (2006) thematic analysis, according to which data were organized in themes or patterns. Nvivo software was chosen to analyze the qualitative data for efficiency and accuracy reasons.

According to Braun and Clarke (2006) the themes are determined through a "rigorous" process including six steps (1) data familiarization, (2) establishing initial codes, (3) identifying themes, (4) reviewing themes, (5) defining and naming themes, and (6) reporting. First, the process of data familiarization was internalized through transcription of the five interviews and the notes (main points and short quotes) from the other interview to Nvivo software. Then, the transcripts of the five interviews were imported into Nvivo, and the coding (nodding) and themes extracting stages began. Next, the major themes of the

Table 4 Descriptive statistics and T-tests of the students' game experiences playing NNG 1 and NNG 2

Dimensions	NNG 1 (N=642) M (SD)	NNG 2 (N=526) M (SD)	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
Flow	2.04 (0.83)	2.47 (0.95)	-7.54	923	0.000	.48
Positive affect	2.28 (1.05)	2.90 (1.1)	-9.34	1029	0.000	.58
Positive value	2.29 (0.95)	2.66 (0.99)	-6.12	1051	0.000	.38
Competence	3.06 (0.96)	3.1 (1.01)	-6.6	1042	0.509	.04
Immersion	2.01 (0.91)	2.51 (1.09)	-7.88	928	0.000	.50
Challenge	2.35 (0.79)	2.63 (0.77)	-5.62	1023	0.000	.36
Negative affect	3.06 (1.01)	2.37 (0.96)	11.1	997	0.000	.70
Tension	2.55 (1.26)	1.81 (0.99)	10.6	1047	0.000	.65

interviews such as “challenges”, “new features” and “positive feelings” were extracted from the data. Quotes by students were noted for further use in the Results and Discussion sections.

Results

Results are presented in two main subsections. First, we describe the students' game experiences playing three versions of the NNG and explore how changes in the NNG development related to gaming experiences. Next, students' preferences and attitudes towards added extrinsic game elements in NNG 4 will be discussed.

Gaming experiences in three versions of NNG

NNG 1 and NNG 2 were played by similar set of participants with students from 4th grade to 6th grade, an independent-samples *t*-test was conducted to compare students' game experiences playing NNG 1 and NNG 2 (Table 4).

Results indicated that students playing NNG 1 rated their game experiences somewhat negatively as Negative Affect ($M=3.06$, $SD=1.01$) was higher than the Positive Affect ($M=2.28$, $SD=1.05$), whereas students playing NNG 2 reported to be having more neutral and positively inclined experience when Positive Affect ($M=2.90$, $SD=1.1$) was higher than Negative Affect ($M=2.37$, $SD=0.96$). Results from an independent *t* test showed that there were significant differences between game experiences in all dimensions, except Competence with effect sizes ranged from close to medium ($d=0.36$) to close to large ($d=0.7$). Tension and Negative Affect decreased by about 0.70, while all other dimensions increased, indicating that improving the usability and clarity of the game reduced negative experiences, and students playing NNG 2 were less likely to feel irritable or annoyed compared to students playing NNG 1. Competence and Challenge of NNG 2 were higher than that of NNG 1, but the difference was not significant in Competence score suggesting that added features (the pirate ship and hidden operations) in NNG 2 were likely to make the game more challenging.

Since NNG 1, NNG 2 and NNG 3 were all played by a set of fourth graders in 2014 and 2015, a one-way ANOVA test was conducted to compare fourth graders' gaming experiences playing three versions of NNG (Table 5).

Table 5 Game experiences of 4th graders playing NNG 1, NNG 2, NNG 3

Dimensions	Game versions			<i>F</i>	<i>p</i>
	NNG 1 (N=63)	NNG 2 (N=72)	NNG 3 (N=40)		
	<i>M</i> (SD)	<i>M</i> (SD)	<i>M</i> (SD)		
Flow	2.23 (0.96)	2.96 (1.07)	2.47 (0.69)	$F(2, 146)=8.43$.000
Positive affect	2.52 (1.20)	3.41 (1.24)	3.18 (1.12)	$F(2, 152)=8.28$.000
Positive value	2.49 (1.09)	3.17 (1.17)	3.0 (1.05)	$F(2, 157)=5.84$.004
Competence	3.04 (0.99)	3.63 (0.96)	3.16 (0.89)	$F(2, 150)=6.14$.003
Immersion	2.1 (1.08)	3.0 (1.19)	2.82 (0.98)	$F(2, 155)=10.62$.000
Negative affect	2.70 (1.01)	2.17 (1.07)	2.07 (0.83)	$F(2, 148)=5.76$.004

Result showed that there were significant differences between game experiences in all 6 dimensions. Post-hoc comparison using Tukey HSD indicated that gaming experiences of NNG1 in all dimensions were significantly different with NNG 2 or NNG 3. This means that overall improvement in later editions of NNG improved students' game experiences in general compared to NNG 1. Between fourth graders' gaming experiences playing NNG 2 and NNG 3, there was no significant differences, except in Flow dimension (Mean difference = -0.49 , $p=0.043$). This means that changes in NNG 3 did not have as positive impacts in game experience as intended.

Students' attitudes and preferences between NNG 3 and NNG 4

Interview results suggested the following significant themes: (a) new extrinsic elements, (b) positive experience, (c) challenges, (d) negative experience, and (e) playing math game vs. doing math exercises.

New extrinsic elements, especially motivating features, were considered the biggest interests for participants. All students showed excitement in having "their own island" (also "the village" or "city page") where they could construct buildings and purchase ships using the materials and coins. Participant 3 stated, "So, when you complete every level, you feel like you own something. You own the building and ships and stuff."

Positive experience could be observed throughout the interviews, such as students stated that the ability to "build" or "own" something made the game more fun and exciting. As NNG 3 and NNG 4 were tested by the same group, students were able to make comparisons of the reward systems between the two versions. Participant 4 noted, "There [in NNG 3], you can only shop the ships, and here, you can build things... I can do things like this (opened Lighthouse Island and started constructing buildings)... also I can buy the big ships."

NNG 4's visual appearance was also considered as a contributing factor in students' positive experience, although they were not mentioned directly. Most participants acknowledged this via praises such as "nice graphics", or comments on the game components such as "the ships are much more beautiful."

General attitudes toward the integration of the in-game tutorial and hint option were quite positive. However, only two out of six participants played the tutorial and found it "helpful" as it showed them "where to go" and "what to do." The rest said they remembered

how to play the NNG. None of the interviewees used hint option during their gameplay. Only one participant “tested” this function. Half of the students did not notice this feature, which could be because they did not play the tutorial where the hint option was introduced. The other half reasoned that they “didn’t need to” or “didn’t want to.” Not wanting or not needing to use the hint function can be related to the Challenge dimension of the game, as five out of six participants reported that the game’s difficulty was important to them and they preferred harder levels. Participant 2 stated, “So, it’s actually fun to have the pirates because it makes the game harder, and you can’t go back using the same way you got there, and you have to plan how to get there... The difficulty makes it more fun, it is not too easy, and it is not just calculation.” Participant 3 said, “The fun part was it [the math] was really challenging, so you had to think when you play. But it was still fun.”

Two of the participants also mentioned that the availability of the hint function might be useful when players advance to more difficult levels. One explained that they did not finish all the maps in NNG 3 so “it was very frustrating.”

Both positive and negative feelings were expressed during the interviews. All participants conveyed positive experiences at some point when discussing their experiences with NNG 4. Adjectives such as “fun,” “good,” and “exciting” referred to the participants’ opinions about the gameplay, while verbs such as “enjoy” and “like” (playing the game) were also used. When asked if they preferred playing the game to doing conventional mathematical exercises, all participants said they preferred playing the NNG.

Regarding negative experiences, students expressed there were some problems with the gameplay that hindered their experiences. Most of the answers were linked to technical issues in NNG 4, mentioned by some participants as “bugs” or “errors.” This was expected as some functions of NNG 4 were still under development. One participant also expressed that to them the game was “boring” and “simple” because “it’s just math.”

Discussion

Despite the strong interest, there is a lack of literature about students’ gaming experiences throughout the development process (De Grove et al. 2010), details accounts concerning different design stages and design choices in making serious games (Habgood 2007; Oksanen 2013; Wouters and van Oostendorp 2013). In this study, we described the developmental process documenting changes between four iterations of the NNG over three year, and rationales behind those decisions. We also investigated students’ gaming experiences playing these versions of NNG to understand how those changes impacted students’ game experiences.

Findings from students of three grade levels playing NNG 1 and NNG 2 showed that gaming experiences were statistically significantly improved in NNG 2. Differences can be observed most clearly in Negative Affect, Tension and Positive Affect dimensions respectively, which implied that students were much less likely to feel irritable, annoyed or frustrated, and more likely to feel good and positive while playing NNG 2. The outcomes can be understood as the changes made in NNG 2 with improvement in the usability of the game mechanism and in-game interface had greatly eased the gameplay. The findings are similar to works of De Grove et al. (2010) about different stages in serious game design, which suggested that improvements in usability could be linked to game experience. Furthermore, the need to repeat many previously played steps in new gaming sessions in the NNG 1 was frustrating for students. During the

play, students in NNG 1 got very little information of their progress whereas NNG 2 gave more clear information about the current game mode and own progress in relation to the criteria. This upholds the argument of Hunicke et al. (2004) on the importance of attention towards game design, and how small changes can have different impacts on the gaming experience and learning process. Added features like the pirate ship and hidden operations could also make the gaming more challenging, and may have had positive impacts on students' game experiences compared to NNG 1. Lastly, the improvement in gaming experiences between students playing NNG 1 and NNG 2 can also be partly explained with the differences in studying designs of NNG 1 and NNG 2. Players used NNG 1 for 10 weeks as part of their regular mathematics classes, whereas players used NNG 2 more intensively during shorter time (about 5 weeks).

Results from fourth graders' gaming experiences playing NNG 1, NNG 2 and NNG 3 indicated that gaming experiences were significantly improved in both NNG 2 and NNG 3 compared to NNG 1. This is expected since NNG 3 contained similar changes to NNG 2 with some additional features such as the level arrangement, reward system and Help page. However, between players' gaming experiences playing NNG 2 and NNG 3, there was no significant improvement, and even decreased in some dimensions such as Flow. This means that while intrinsically the game mechanics were similar, changes in the level arrangement, addition of the reward system (Shop where it was possible to purchase more impressive ships with the earned coins) and Help page in NNG 3 did not have the intended positive effects on students' gaming experiences. When compared with commercial entertainment games, these rewarding features were modest. New features of the NNG 2 that improved the usability of the original game version (NNG1) seemed to be more important for students' game experience than the additional modest reward and help functions of the NNG 3.

Findings from interviews also shed some lights on players' gaming experiences in NNG 3 as participants were able to compare NNG 3 and NNG 4. As some participants recalled, in NNG 3, even though players could purchase ships at the Shop page, they did not feel like they "own" the ships. However, in NNG 4, having a separate island designated for constructing buildings and ships, players were able to create their own game worlds with their accomplishments (materials and coins) from the maps. Interviewed students attributed the feeling of ownership in NNG 4 to the fact that they could see the transformation happening gradually, and seeing the transformation or "building" something themselves made the game itself more exciting or interesting.

This deserves some discussion, as it could be very beneficial for future serious game developments. Since the motivating elements in NNG 3 and NNG 4 virtually served the same purpose, which is to give meaning to the coins and material earned during gameplay, the way they were designed clearly set them apart. Therefore, it is possible to compare between the extrinsic designs changes in these two versions. In NNG 3 (and previous versions), it was implied that those materials were used for building the villages as a pop-up window displayed an image of a village changing from basic to more modern, and coins earned were used for purchasing ships. However, not taking part in that transformation seemingly did not give players the feeling of ownership as it did while playing NNG 4.

From a game design perspective, this indicates that it is important to pay attention to how the reward system and/or motivating mechanism is designed. The idea of giving meaning to the coins and materials, or using the coins to purchase the Ships is similar between two versions; however, not being able to see the transformation of the Ship or the City seems to take away some joy, as the students did not feel like they "own" the ships or

the city. From a technical point of view, the transformation is only possible with the Unity Game Engine and three-dimensional game, whereas it was not as exciting as it was in the two-dimensional graphic design.

These results show that apart from aiming to integrate the educational content and basic game features (Habgood and Ainsworth 2011a, b) it is also important to pay attention to the integration of meaningful use of extrinsic game elements (such as reward systems and game story) as this may help to maintain players' interest and motivation in playing the game. This is also in alignment with arguments about reward systems' benefits that have been explored extensively in digital games. Hallford and Hallford (2001) believed that well-designed reward systems could prolong players' excitement during gameplay and maintain positive gaming experiences and motivations.

In previous versions, the game's interfaces had been gradually improved with better visual feedback cues (NNG 2, NNG 3) which resulted in improved Immersion scores; however, the graphic design and sound was not prioritized. This issue was addressed in the enhancement of the visual aesthetics of NNG 4, with more purposefully graphic design and the addition of sounds. Results from interviews also confirmed that students appreciated better game graphic and design in NNG 4. This supports Schell's arguments (2014) that good aesthetics can make up for game design imperfections and increase game attractiveness.

Help functions such as tutorials and hint options were also deemed useful. Students recognized the usefulness of the hint function as they recalled their frustration at not being able to finish all the maps in NNG 3, and believed that this option would have been helpful in such situations. In addition, all participants acknowledged the more frequent appearance of the pirate ship. As discussed above, the pirate ship was first added in NNG 2 and appeared only during some of the last maps. However, in NNG 4, the pirate ship appeared almost immediately and more frequently. Not only did the pirate ship make the math harder, but when or where it showed up was also unexpected and unknown making the game more exciting and challenging than usual. This is relevant in both learning and game design, as mathematic training requires variation in number-operation use and the pirate ship does exactly that.

Limitations and future studies

First, it was not possible to have participants from all grade levels from fourth to six to play all versions of the NNGs. Having only fourth graders' gaming experiences playing NNG 1 to NNG 3 also has some limitations, and the result cannot be generalized to all players at all intended ages. Also, all data were based on subjective self-reports resulted in limited understanding of students' game experiences. Future studies could explore the roles of teachers and focus more on different types of data (journals, log files, observations), especially regarding the implementation and teaching practices with NNG in the classrooms.

Second, during NNG 4 study, the game was not fully developed, not all maps were available, and the scoring procedure was still under construction. Thus, the experiment was short-lived. As participants were familiar with the core game mechanics, some features in NNG 4 were not fully tested as participants overlooked those functions (such as Tutorial and Hint function). Additionally, there was no novelty effect on participants because NNG 4 technically was not a new game to them. Different results might be achieved if participants with no previous experience played the NNG for longer period. Additionally, only a

small number of participants took part in the interview, which resulted in some limitation in understanding students' experiences in general.

Conclusions and implications

Results indicate that improvement in game's usability and clarity in the user interface were effective in providing more positive, smooth and immersive game experiences for players. Therefore, serious game designers should put effort and focus on these aspects of the game, in addition to focusing on the educational content. Further work is also needed to investigate the exact value that extrinsic elements could add to maintain players' enthusiasm and situational interest in serious games, as not all motivating elements have the intended impacts on students' gaming experiences. Furthermore, there seems to be a clear advantage in having better aesthetics in game-based learning as players preferred and appreciated the nicer design and graphic. Above all, the outcome of this study is beneficial for game developers and researchers. The study provides a comprehensive description and analysis of the developmental process of making a serious game. This can serve as an example to determine which features are needed to create meaningful serious games and how changes in game design influence students' game experiences.

Acknowledgements Open access funding provided by University of Turku (UTU) including Turku University Central Hospital.

Funding This work was supported by the Academy of Finland [Grants Number 274163 and 312528; PI: Erno Lehtinen] and the Academy of Finland DECIN & MAY projects [Grant Number 278579 and 305619; PI: Minna Hannula-Sormunen].

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in this study were following strictly the ethical guidelines of the University of Turku, Finland.

Informed consent Participation was voluntary and informed consent was obtained in writing from all the participants' parents.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

Structure of the Game Experience Questionnaire with Statements in Each Dimension (Scales range from 1 to 5; 1 = Not at all, 5 = Extremely).

Dimensions	Statements
Flow	I felt completely absorbed by the game I forgot everything around me when I played I lost track of time when I played I was deeply concentrated in the game I was fully occupied with the game
Positive affect	I felt content when I played I felt happy when I played I felt good when I played I enjoyed playing I thought playing was fun
Positive value	I have gotten better at math after playing I like math more after playing this game This game helped me learn math
Competence	I felt skillful when I played I was good at playing I felt successful when I played I felt I was good at the game
Immersion	I was interested in the game's story I felt imaginative when I played I felt I could explore how to navigate around numbers when I played
Challenge	I felt that I was learning when I played I thought playing this game was hard I thought the game was difficult enough I had to put a lot of effort when I played
Negative affect	I was thinking about other things while playing I thought playing was boring I was bored while playing The game's idea felt boring Playing gave me a bad mood
Tension	Playing annoyed me I felt irritable when I played

References

- Abeele, V. V., De Schutter, B., Geurts, L., Desmet, S., Wauters, J., Husson, J., et al. (2012). P-III: a player-centered, iterative, interdisciplinary and integrated framework for serious game design and development. In S. De Wannemacker, S. Vandercruyse, & G. Clarebout (Eds.), *Serious games: The challenge* (pp. 82–86). Berlin: Springer.
- Birk, M. & Mandryk, R. L. (2013). *Control your game-self: Effects of controller type on enjoyment, motivation, and personality in game: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 685–694). New York: ACM. <https://doi.org/10.1145/2470654.2470752>.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 0887, 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- Brezovszky, B. (2019). *Using game-based learning to enhance adaptive number knowledge*. University of Turku. Retrieved May, 2019, from <https://www.utupub.fi/handle/10024/147025>
- Brezovszky, B., Lehtinen, E., McMullen, J., Rodriguez, G. & Veermans, K. (2013). *Training flexible and adaptive arithmetic problem solving skills through exploration with numbers: The development of number navigation game. European Conference on Games Based Learning*. Reading: Academic Conferences International Limited. Retrieved May, 2019, from <https://search.proquest.com/docview/1549957336?accountid=14774>.

- Brezovszky, B., McMullen, J., Veermans, K., Hannula-Sormunen, M. M., Rodríguez-Aflecht, G., Pongsakdi, N., et al. (2019). Effects of a mathematics game-based learning environment on primary school students' adaptive number knowledge. *Computers & Education*, *128*, 63–74. <https://doi.org/10.1016/J.COMPEDU.2018.09.011>.
- Brezovszky, B., Rodríguez-Aflecht, G., McMullen, J., Veermans, K., Pongsakdi, N., Hannula-Sormunen, M. M., et al. (2015). Developing adaptive number knowledge with the Number Navigation game-based learning environment. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 155–170). Cham: Springer.
- Brown, E. & Cairns, P. (2004). *A grounded investigation of game immersion: CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1297–1300). New York, NY, USA: ACM. <https://doi.org/10.1145/985921.986048>.
- Bruckman, A. (1999). *Can educational be fun? Game Developer's Conference*.
- Burnes, D., Butchko, J., & Patrick, S. (2015). Maximizing effectiveness of educational games through gaming experience. *From Andrewd. Ces*. Retrieved September, 2018, from <https://andrewd.ces.clemsu.edu/courses/cpsc414/spring14/papers/group5.pdf>.
- Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. New York: Harper & Row.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety. Beyond boredom and anxiety*. San Francisco: Jossey-Bass.
- De Grove, F., Van Looy, J., Courtois, C. & De Marez, L. (2010). "I play, therefore I learn?" *Measuring the Evolution of Perceived Learning and Game Experience in the Design Flow of a Serious Game: Proceedings of meaningful play*. (pp. 1–32). East Lansing: Michigan State University.
- Eck, R. V. (2006). Digital game-based learning : It's not just the digital natives who are restless *Educator Review*, *41*(2), 1–16. <https://doi.org/10.1145/950566.950596>.
- Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, *32*(3), 158–172. <https://doi.org/10.1007/s11031-008-9102-4>.
- Ermí, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. *Changing Views: Worlds in Play*. <https://doi.org/10.1080/10641260490479818>.
- Gaydos, M. (2015). Seriously considering design in educational games. *Educational Researcher*, *44*(9), 478–483. <https://doi.org/10.3102/0013189X15621307>.
- Habgood, M. J., & Ainsworth, S. E. (2011a). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, *20*(2), 169–206.
- Habgood, M. J., & Ainsworth, S. E. (2011b). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *The Journal of the Learning Sciences*, *20*(2), 169–206. <https://doi.org/10.1080/10508406.2010.508029>.
- Habgood, M. P. (2007). *The effective integration of digital games and learning content*. University of Nottingham.
- Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., & Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education*, *102*, 202–223. <https://doi.org/10.1016/J.COMPEDU.2016.09.001>.
- Hallford, N., & Hallford, J. (2001). *Swords and Circuitry: A designer's guide to computer role playing games*. Roseville, CA: Prime Publishing.
- Hamari, J., & Kehonen, L. (2017). Why do people play games? A meta-analysis. *International Journal of Information Management*, *37*(3), 125.
- Hunicke, R., LeBlanc, M. & Zubeck, R. (2004). *MDA: A formal approach to game design and game research: Proceedings of the AAAI Workshop on Challenges in Game AI*.
- Jaatinen, M. (2016). *Motivaatio ja matematiikan oppimistulokset peliympäristössä*. University of Turku. Retrieved May, 2019, from <https://www.utupub.fi/handle/10024/122920%09>.
- Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., et al. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, *66*(9), 641–661. <https://doi.org/10.1016/j.ijhcs.2008.04.004>.
- Ke, F. (2016). Designing and integrating purposeful learning in game play: A systematic review. *Educational Technology Research and Development*, *64*(2), 219–244. <https://doi.org/10.1007/s11423-015-9418-1>.
- Kiili, K., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science*, *15*, 78–91. <https://doi.org/10.1016/J.PROCS.2012.10.060>.
- Kiili, K., & Lainema, T. (2008). Foundation for measuring engagement in educational games. *Journal of Interactive Learning Research*, *19*(3), 469–488. <https://doi.org/10.1016/j.iheduc.2004.12.001>.
- Kiili, K., Lainema, T., Freitas, S. D., & Arnab, S. (2014). Flow framework for analyzing the quality of educational games. *Entertainment Computing*, *5*(4), 367–377. <https://doi.org/10.1016/j.entcom.2014.08.002>.

- Lehtinen, E., Brezovszky, B., Rodríguez-Aflech, G., Lehtinen, H., Hannula-Sormunen, M. M., McMullen, J., et al. (2015). Number navigation game (NNG): Design principles and game description. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 45–61). Cham: Springer.
- Ma, M., Bale, K. & Rea, P. (2012). *Constructionist learning in anatomy education: What anatomy students can learn through serious games: Proceedings of the Third International Conference on Serious Games Development and Applications* (pp. 43–58). Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-642-33687-4_4.
- McMullen, J., Brezovszky, B., Hannula-Sormunen, M. M., Veermans, K., Rodríguez-Aflecht, G., Pongsakdi, N., et al. (2017). Adaptive number knowledge and its relation to arithmetic and pre-algebra knowledge. *Learning and Instruction, 49*, 178–187. <https://doi.org/10.1016/J.LEARNINSTRUC.2017.02.001>.
- Mitgutsch, K. & Alvarado, N. (2012). *Purposeful by design?: A serious game design assessment framework: Proceedings of the International Conference on the Foundations of Digital Games* (pp. 121–128). New York, NY, USA: ACM. <https://doi.org/10.1145/2282338.2282364>.
- Oksanen, K. (2013). Subjective experience and sociability in a collaborative serious game. *Simulation & Gaming, 44*(6), 767–793. <https://doi.org/10.1177/1046878113513079>.
- Poels, K., de Kort, Y. & Ijsselstein, W. (2007). FUGA: The fun of gaming: Measuring the human experience of media enjoyment. Deliverable D3.3: Game Experience Questionnaire.
- Rodríguez-Aflecht, G., Jaakkola, T., Pongsakdi, N., Hannula-Sormunen, M., Brezovszky, B., & Lehtinen, E. (2018). The development of situational interest during a digital mathematics game. *Journal of Computer Assisted Learning, 34*(3), 259–268. <https://doi.org/10.1111/jcal.12239>.
- Rodríguez-Aflecht, G., Brezovszky, B., Pongsakdi, N., Jaakkola, T., Hannula-Sormunen, M. M., McMullen, J., et al. (2015). Number navigation game (NNG): Experience and motivational effects. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 171–189). Cham: Springer.
- Schell, J. (2014). *The art of game design: A book of lenses* (2nd ed.). Natick, MA: A. K. Peters Ltd.
- Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computers in Entertainment, 3*(3), 3–3. <https://doi.org/10.1145/1077246.1077253>.
- Tashakkori, A., & Teddlie, C. (2010). *SAGE handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781506335193>.
- Torbeyns, J., Lehtinen, E., & Elen, J. (2015). Describing and studying domain-specific serious games. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 1–250). Cham: Springer.
- Weibel, D., & Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology, 10*. <https://doi.org/10.1155/2011/282345>.
- Whitton, N. (2011). Game engagement theory and adult learning. *Simulation & Gaming, 42*(5), 596–609. <https://doi.org/10.1177/1046878110378587>.
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249–265. <https://doi.org/10.1037/a0031311>.
- Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education, 60*(1), 412–425. <https://doi.org/10.1016/J.COMPE DU.2012.07.018>.
- Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., et al. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research, 82*(1), 61–89. <https://doi.org/10.3102/0034654312436980>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Phuong Bui is working as a Ph.D. candidate at Centre for Research on Learning and Instruction and Department of Teacher Education, University of Turku, Finland.

Gabriela Rodríguez-Aflecht is working as a postdoctoral researcher at Centre for Research on Learning and Instruction and Department of Teacher Education, University of Turku, Finland.

Boglárka Brezovszky is working as a postdoctoral researcher at Centre for Research on Learning and Instruction and Department of Teacher Education, University of Turku, Finland.

Minna M. Hannula-Sormunen is a Professor at Centre for Research on Learning and Instruction, Department of Teacher Education University of Turku, Finland.

Samuli Laato is working as a Project Researcher at Department of Educational Sciences, and a Ph.D. candidate at Department of Future Technologies, University of Turku, Finland.

Erno Lehtinen is a Professor at Centre for Research on Learning and Instruction and Department of Teacher Education, University of Turku, Finland.