Flow Experience as a Quality Measure in Evaluating Physically Activating Collaborative Serious Games

Kristian Kiili, Arttu Perttula, Antero Lindstedt, Sylester Arnab, Marko Suominen

Flow Experience as a Quality Measure in Evaluating Physically Activating Collaborative Serious Games

Kristian Kiili[1], Arttu Perttula[1], Antero Lindstedt[2], Sylester Arnab[3,] Marko Suominen[1]
[1]*Tampere University of Technology, {firstname.lastname}@tut.fi*
[*2]*Flow Factory Ltd, antero.lindstedt@flowfactory.fi*
[3]*Author Disruptive Media Learning Lab, s.arnab@coventry.ac.uk*

Abstract

The measurement of subjective playing experience is an important part of a game development process. The enjoyment level that a serious game offers is a key factor in determining whether a player will be engaged in a gameplay and achieve the specific objectives of the game, such as learning. In this paper, we report the results of a game design process in which two prototypes of a collaborative physically activating games were studied. The main aim of the paper is to explore to what extent the measurement of flow experience can facilitate game evaluation and design process. Altogether 102 junior high school students participated in two user experience studies and played physically activating collaborative game designed to teach soft skills. Playing experience was measured using flow questionnaire, playing behaviour was observed and some of the players were interviewed. Together, the results showed that flow experience can be used to evaluate the overall quality of the gameplay and it provides a structured approach to consider the quality of the game. However, flow does not provide detailed information about the weaknesses of the game and thus complementary methods are necessary to identify the causes. The results also indicated that flow experience was independent of gender, which supports its use in quality measurement.

Keywords: *Game design, Flow experience, Exergame, Quality measurement, Collaboration*

1. Introduction

The evaluation of subjective playing experience is a crucial part of a game development process. The enjoyment level that a serious game offers is a key factor in determining whether a player will be engaged in the gameplay and achieve the desired objectives of the game, such as learning. Thus, the ability to quantify the playing experience is an important goal for both industry and academia. In general, game developers require a reliable way to measure the overall engagement level of their games and to pinpoint specific areas of the experience that need improvement [1]. Several constructs have been proposed to describe playing experience, but definitional agreement has not been achieved [2]. The most common concepts that have been linked to playing experience are engagement, involvement, immersion, presence, motivation and flow. According to Procci, Singer, Levy and Bowers [1] the concept of flow is one of the most popular constructs that describe gaming experience. For example, several authors [3-7] have considered playing experience in terms of flow. The research has shown that flow reflects the enjoyment that game playing produces [e.g. 8] and it has a positive influence on motivation and learning [8, 9]. However, the empirical research regarding the relation between flow, enjoyment, and learning is argued to be limited and more empirical research is needed [8].

In this paper, we consider the usefulness of flow experience in evaluating the quality of physically stimulating serious games (exergames). We report the results of the game design process in which a collaborative exergame aimed at teaching soft skills was developed. In practice, two versions of the game, namely Labyrinth Run (first iteration, n = 53) and TeamBoost (second iteration, n = 49) were studied in terms of flow experience. In order to understand factors that affect the playing experience of collaborative exergame, we next discuss the flow construct, design principles for educational exergames and characteristics of cooperative (collaborative) games. With these perspectives, the details and the results of the two studies are presented and finally the usefulness of flow construct in evaluating the playing experience is discussed.

2. Flow experience

Flow describes a state of complete absorption or engagement in an activity and refers to the optimal experience [10]. During the optimal experience, a person is in a psychological state, where he or she is so involved with the goal-driven activity that nothing else seems to matter. An activity that produces such experiences is so satisfying or pleasurable that the person may be intrinsically motivated to engage, without being concerned with what he will get out of his action. This kind of motivation is very important especially in serious games that usually require different cognitive or physical investments compared to entertainment games. Czikszentmihalyi's [10] flow theory subsequently has been applied in several different domains including, for example sports, art, work, human–computer interactions, games and education. In the area of games, it has a particular value, as it maps well against the intrinsically motivating activities [3]. Theoretically, flow consists of nine dimensions [3,10] including clear goals, challenge-skill balance, immediate feedback, sense of control, action awareness merging (some times operationalized as playability in serious games context [11-13]), loss of self-consciousness, concentration, autotelic (rewarding) experience, and time distortion.

Goal setting is a process that, when undertaken correctly, helps to move a player towards flow. When a player is in flow, the clarity of purpose occurs on a moment-by-moment basis and focus on a task and response to cues that a game provides can be optimised. If the goals seem too challenging or are hard to perceive, the probability of experiencing flow is low. If the objectives of a serious game are discrete from gameplay, the game may fail to produce educationally effective experiences.

Paying attention and interpreting the *feedback* that the game provides is important in evaluating whether one is on a track towards achieving the desired goals or not and helps in monitoring playing performance. The feedback dimension can be divided into immediate and unambiguous feedback as well as into cognitive feedback. The immediate and unambiguous feedback keeps a player focused and helps him/her to perceive the consequences of his/her actions. If a player has to wait long before he can realize what effect his action caused, he will become distracted and loose the focus on the task, progress against goals and the events of the game. Additionally, the delayed or confusing feedback may create interpretation problems and even lead to misconceptions and negative learning transfer. The cognitive feedback relates to the cognitive problem solving and provides the account for learning. The cognitive feedback aims to stimulate a player to reflect on his experiences and test solutions in order to further develop his mental models [14] and playing strategies. In other words, it focuses a player's attention on information that is relevant for the learning process.

In this paper, the *playability* is included to partly replace Csikszentmihalyi's [10] *action-awareness merging* dimension, which is problematic in the educational game context. According to Csikszentmihalyi, all flow inducing activities become spontaneous and automatic, which is not desirable from a learning point of view. In contrast, the principles of experiential and constructive learning approaches gives emphasis to the fact that learning is an active and conscious knowledge-construction process. It is noteworthy that reflection is not always a conscious action by a player. However, only when a player consciously processes his experiences can he make active and aware decisions about his playing strategies and thereby form a constructive hypothesis to test. Thus, a distinction between activities related to solving problems or learning and controlling the game should be made. This means that controlling the game should be spontaneous and automatic, but the educational content related to a player's tasks should be consciously processed and reflected.

Generally, the aim of a game design targeting learning objectives is to provide students or learners with *challenges* that are balanced with their skill level. If the challenge is too low, a player tends to experience boredom and when the challenge is too high, a player tends to feel anxiety. Furthermore, challenges should be related to the main task so that the flow experience is possible. When both the task and the use of the artefact are complex, then the artefact and the task may detract from the player's attention. In fact, bad playability decreases the likelihood of experiencing task-based flow because the player has to sacrifice attention and other cognitive resources to the inappropriate activity. Because the information processing capacity of working memory is limited [15], all possible resources should be available for relevant information processing (the main task) rather than for the use of the game controls (if the learning objective is not bound to game controls). Thus, the aim of the user interface design of games is to support the shift from cognitive interaction to fluent interaction. In an ideal situation, the controls of the game are transparent and allow the player to focus on higher order cognition rather than solely upon controlling tasks.

Sense of control clearly relates to the challenge-skill balance dimension. Csikszentmihalyi [10] has stated that sense of control refers to possibility rather than to actuality of the control. It can be said that a person senses when he can develop sufficient skills to reduce the margin of error close to zero, which makes the experience enjoyable. The feeling of being in control frees a player from thoughts of failure and thus he/she is encouraged to perform more creatively and exploratively. Total concentration is one of the most powerful feelings of flow. Because flow-inducing activities require complete *concentration* of attention on the task at hand, there are no cognitive recourses left over for irrelevant information. Such concentration provides great satisfaction, which in turn leads to growth in complexity. When in flow, a player does not have to invest effort to keep his mind on the task.

Most of us live in the world of evaluation. During important activities such as learning, it is hard to stop thinking how others evaluate us. When a player can ignore what others think of him or her, the player has *lost self-consciousness*. The problem is that the criticism that the player may face turns his attention away from the actual task and turns too much to self, which does not facilitate the performance and playing experience. Self seems to disappear from awareness during flow and thus in flow there is no room for self-scrutiny [10]. Here, self refers to the self-esteem and thus loss of self-consciousness does not limit reflective thinking processes.

High flow feelings tend to transform players' *perception of time*. According to Csikszentmihalyi [10], the sense of time tends to bear little relation to the passage of time as measured by the absolute convention of a clock during the flow experience. Time seems to either pass really fast or the seconds may feel like minutes.

Autotelic or rewarding experience refers to an activity that is done, not with the expectation of some future external benefit, but simply because the doing itself is interesting and fun. The experience is so satisfying and enjoyable that player is motivated to experience it again. In fact, Csikszentmihalyi [10] described this autotelic experience as an end result of other dimensions that provides high motivation towards further involvement with the activity. Although winning is important in many games, flow does not depend on the final outcomes of an activity, and offers players something more than just a successful outcome. In fact, an optimal experience usually occurs when a person's body or mind is stretched to its limits in a voluntary effort to accomplish something difficult and worthwhile [10]. Such experiences are not necessarily pleasant when they occur, but they still produce enjoyment.

3. Design principles for educational exergames

Kiili and Perttula [16] proposed the concept of educational exergames. Educational exergames combine gameplay elements from educational games (cognitively challenging games involving an educational goal) and exergames (physically challenging games). Educational exergames are an emerging form of computer games that aim to leverage the advantages of sports and cognitive training in order to foster physical, social and mental health benefits. In general, educational exergames are a new and unstudied branch of research in the era of serious games. Kiili and Perttula [16] proposed a framework for designing educational exergames. The framework for Exertion games [17], Dual flow model [6], Framework for sports engagement [18], Persuasive technology [19], and Cognitive load theory [20] form the foundation of the Educational exergame framework. The aim of the framework is to provide theoretical means to balance the amount of physical, cognitive, and sensomotoric workloads in order to optimize learning and health effects as well as to describe ways to create more engaging exertion and learning experiences mediated by technology.

In educational exergames, the body and mind play the central role and in the framework, those are used as lenses to consider other aspects of educational exergames. When the rules, game elements and playing context are considered through exertion and learning interaction lenses, several focus points that are crucial for designing educational exergames can be identified. Focus points highlight different aspects of game elements and provide conceptual guidance for design and analysis. From a total of ten focus points concerning the Flow experience, awareness of complexity', and rhythm and intensity of exertion and learning are most essentials for the study reported in this paper and they are briefly discussed below (see [16] for more details).

3.1. Flow in educational exergames

The dimensions of flow experience were already discussed above, but challenge-skill balance dimension needs more clarification in the educational exergame context. According to the extended dual flow model [6], challenge-skill balance needs to be considered from both cognitive and physical perspectives. In the model, the intensity-fitness balance determines the effectiveness of an exertion. If the game is too intensive, a player will fail to play the game and is unable to continue exercising. On the other hand, if the intensity is too low compared to a player's fitness level, he/she will enter a state of deterioration. The optimal exergaming experience can be achieved when both the psychological (cognitive) and physiological challenges are in balance and a player is in the flow zone. In general, the challenge of a game design is to keep the player in a flow state by balancing both the cognitive and physiological aspects of the game.

It is evident that the balancing of cognitively demanding exergames is not as straightforward as the balancing of traditional computer games. The basic balancing principle suggests that the difficulty level of a game can be gradually increased because it is assumed that players' skill level increases with playing time. Such an approach does not work properly in cognitively demanding exergames. Even though a player's skills may increase during playing, the gradually increasing intensity will finally lead to exhaustion and failure in lengthy playing sessions. To overcome this problem, Sinclair, Hingston & Masek [6] suggested that exergames should adapt dynamically to a player's performance, or they should be based on simple mechanics that focus more on input devices and exercise movements than on complex game play.

3.2. Awareness of complexity

In general, complexity can be defined as a state of the system that involves numerous elements and numerous relationships among these elements. In educational exergames, the complexity is composed from several factors, such as the amount of bodily controls, the amount of simultaneous players, the type and the level of learning content, the audio-visual implementation of the game, the rules of the game, and of course the relationships between these factors. Players should be able to form clear conceptual models of the game that direct their behaviour in the game.

When designing educational exergames, we should remember the constraints of human cognition and thus design the gameplay according to the target group's skills, characteristics and knowledge. When playing educational exergames, learners are challenged to extract relevant information from a game world, select corresponding parts of information and integrate all of these elements to form coherent representation and at the same time track the state of the game, decide the right movements to carry out, move, possibly communicate with other players, and interpret bodily sensations. This requires a lot from the player. Due to the nature of a game world that changes during playing, important information may be presented only for a while. Therefore, it requires information to be kept actively in working memory in order to integrate it to earlier presented information and relate it to one's actions. This may easily impose too high of a cognitive load in learners cognitive system and hinder learning and playing. Thus, the designers should be aware about the level of complexity and optimize it based on player characteristics.

3.3. Rhythm and Intensity of Exertion and Learning

The dual flow model argues that the physical and cognitive aspects of an exergame need to be balanced. In educational exergames the balancing is even more important and challenging. The rhythm of gameplay and intensity of physical activities plays a central role in this. According to Tenenbaum [21], exercise intensity impacts the focus of attention. Thus, the integration of learning contents and exergame interfaces raises new design challenges. Research on sports has shown that when a physical workload increases, attention allocation shifts from dissociation to association [22]. Association can be defined as turning the focus inward and towards bodily sensations, while dissociation is focusing outward and away from body sensations [23].

Such change in natural attention disturbs the processing of game elements and may also hinder learning and problem solving. In other words, during high physical workload, it is hard to concentrate on problem-solving and game stimuli designed to enhance learning. In general, we can say that the higher the sum of cognitive and physical workloads is, the higher the possibility of failure in the game is as well. For example, the recent study [24] showed that when the sum of cognitive and physical workload was high, players had difficulties to either solve the cognitive tasks or control their game characters properly. The balancing of workloads and adaptation to players characteristics is very challenging because the cognitive and physical workloads are composed of several factors as discussed as part of Complexity focus point. One solution in order to avoid cognitive or physical overload is to sequence the cognitive gameplay and physical gameplay. For example, players could first conceptually solve the puzzle and then perform the solution by controlling the game with physical movements.

4. Competition, cooperation and collaboration in games

Azadegan and Harteveld [25] emphasised that even though games are known for being competitive, collaboration is an essential part of many games, where players combine their efforts to manage a challenge that is too difficult to be dealt with by one individual player. Zagal et al. [26] elaborate on the difference between cooperative and collaborative games, which we aim to briefly discuss to set the context for the case studies explored in this paper.

Collaboration and cooperation are two distinct and complex concepts [27]. According to Zagal et al. [26], cooperative games are based on opportunities that enable players to achieve a win-win condition, emphasising on participation, challenge, and fun rather than competition. In collaborative games on the other hand, players form a team and share pay-offs and outcomes. If a team wins, everyone wins. If a team loses, everyone loses. Collaborative gameplay is a feature that allows players to work together as teammates against one or more opponents. Therefore, there is a level of competition implemented to drive the dynamic of the overall game ecosystem. In some multiplayer games, the opponents could be non-player characters (AI teams). In this case, the game is designed in such a way that players are playing against the machine, which promotes cooperation and collaboration between the players. Playing simultaneously allows players to assist one another and create playing strategies as a team.

Collaborative as well as cooperative games are usually very motivating and external pressure or rewards are not needed. Based on Malone's and Lepper's [28], group-level intrinsic motivators that involve interaction among people, cooperation is a motivator, one that seems to be built into human nature. When people belong to a team, most of them cooperate and collaborate. And in terms of collaborative learning as promoted by such a serious game, Brufee [29] elaborated that the fundamental premise of collaborative learning is based upon consensus building through cooperation among group or team members. Romero et al. [30] stated that serious games could promote intragroup cooperation and intergroup competition. Games such as the multiplayer Course sans Gagnant game [31] encourage intra-team collaboration and inter-team competition. Based on a car racing game, the winning team will be the team with players reaching the finish line at the same time. Intra-cooperation is promoted within the team, where players share strategies and maths knowledge in order to calculate the different speeds required to finish the race together.

Cooperative and collaborative gameplays therefore include overlapping features. Sedano et al. [27] emphasise that although there is a distinction, the terms are often used interchangeably. For the purpose of our case studies, we define both cooperative and collaborative gameplay as the activity in which players come together to achieve a single goal, and competition can be used as a mechanic to drive the group dynamics. The cooperative, collaborative and competitive characteristics can be used to ensure motivation is sustained and contextual immersion is achieved which will lead to the desired flow in the play-learn experience.

To aid the design of cooperative games, Rocha, Mascarenhas and Prada [32] described six Design Patterns for Cooperative Games that were applied in the games used in this research. These patterns are:

- ***Complementarity*** is **o**ne of the most commonly used design patterns in cooperative games. The use of the pattern ensures that there is some complementarity between the characters that players control. Complementarity between the characters usually leads to several consequences. First, characters tend to settle better in one type of role. Second, even if there are two different character types for the same role, they will usually be complementary to one another, as they will have different abilities that will complement each other in that role.
- ***Synergies between abilities*** pattern tries to ensure that some of the abilities of one character type have some synergy with abilities of another character type.
- ***Abilities that can only be used on another player*.** Game can include players with abilities that can only be used on another player. The main purpose of these abilities is to encourage cooperation between players.

- ***Shared Goals*** is a simple design pattern that aims to force players to work together. In practice, a group of players have one non-exclusive goal that can be completed only as a group.
- ***Synergies between goals*** can be used to force players to cooperate together, if they otherwise have different goals.
- ***Special rules for players of the same team*** *pattern* can be used to promote and facilitate cooperation**.** Special rules for players of the same team can be defined for example in a way that an action will have a different effect when committed on a friendly player compared to an unfriendly player.

5. Method and results of the studies

The overall research approach is based on design science [33]. The research process is cyclical in nature, involving planning, taking action, observing, evaluating and reflecting. In this paper we report the evaluation results of two iterations of a collaborative exergame. In the design of the studied games, the design framework for educational exergames [16], design patterns for cooperative games [32] and movement-based game guidelines [34] were utilized.
The research has two main aims: 1) develop a good collaborative and physically activating game and 2) evaluate the usefulness of flow theory as a quality measure. The assumption of the study is that the level of flow experience reflects player's satisfaction in the game and can reveal weaknesses of the game design.
Until now several different methods have been used to study flow experience from which self-reporting techniques have been the most common methods [35]. Thus, the data related to flow was gathered with an 18-item questionnaire developed by the authors. A 6-point Likert-type response format was used. The items included were derived from the GameFlow [36] and FSS-2 questionnaires [37]. The dimensions included were challenge, goal, feedback, playability, concentration, time distortion, rewarding experience, loss of self-consciousness, and sense of control. Each dimension was measured with a scenario-based statement in order to avoid interpretation problems that have appeared in earlier studies. For example, the feedback dimension was operationalized as follows: "The game provided me such a feedback that I was aware how I was performing. I could really perceive the consequences of my actions." The reliability of the used flow questionnaire ($N = 102$) indicates that the flow construct is internally consistent ($\alpha = .91$), which means that all nine dimensions measured the same phenomenon. Next subsections report the results of two iterations of the game design process.

5.1. Study 1: Labyrinth Run

The aims of the Labyrinth Run game are: 1) teach players to work as a team, 2) teach players to take others into consideration, and 3) teach communication and leadership. The overall goal of the game is to complete a mission as a team of players instead of as an individual player, but it also provides scores for each individual, thus enables competition. In order to facilitate collaboration between players, we applied shared goals, synergies between goals, and complementarity design patterns for cooperative games (see section 4) proposed by Rocha, Mascarenhas, and Prada [32].

5.1.1. Description of Labyrinth Run

Labyrinth Run (Fig. 1) is a side view platform game for 3-5 players. There are no opponents in the game, but players compete against the system as a team. Thus, players have shared goals and they are forced to work together [32]. The task of the players is to solve puzzle-like levels in a given time to proceed. Players must cooperate with each other to pick up a key and exit the level from a door that opens after the key is picked up. Players will be awarded points according to how quickly and how many of the players have completed the level before the time runs out. When the time runs out, the players move on to the next level, regardless of how much progress has been made. The prototype version that was used in this research had five levels.

Figure 1. Screenshot of the Labyrinth Run game (zoomed in); one player runs on a treadmill to help another player overstep the gap with the moving platform.

According to Pasch et al. [38], natural control and mimicry of movements influence immersion in movement-based interaction. Thus we tried to implement natural gestures for game controls and display player's movements in the game screen. In practice, a player has a mobile phone in his hand or in the pocket that interprets the light movement into walking and sudden movement into a jump (Fig. 2). When a player walks in place, the game character walks forward on the screen and when the player jumps, the game character will also jump. We did not manage to implement turning of a game character with decent physical movement and thus we had to use substitutive method as follows. When the game character hits an obstacle, it changes the direction of walking.
Based on the *synergies between goals* design pattern [32], there are synchronized goals, obstacles, in the game that require players to co-operate and work together. Simultaneous game play allows players to assist one another by performing cooperative manoeuvres, such as boosting a teammate up and over obstacles. For example, the sliding door will open only when one of the game characters is standing on the weight trigger. When a game character runs on the treadmill, elsewhere in the game is a horizontally moving platform, which can carry other game characters. One of the levels has two platforms that rise and fall like scales, depending on how many and how heavy game characters are on them. As *complementarity* design pattern suggests, players play different character roles that complement each other [32]. In Labyrinth Run, this means that some of the characters in the game are sturdier and weigh more, which influence their ability to move. Players must agree on every level what each one of them does. Labyrinth game also involves a competition aspect. Each player is awarded with individual points according to their performance. As mentioned previously, cooperation is necessary to solve the challenging levels.

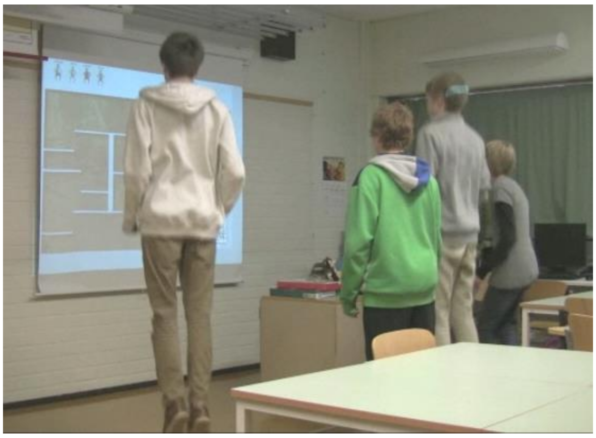

Figure 2. Players move their game characters by moving themselves.

5.1.2. Participants and procedure

The study was conducted in autumn 2012 at a Finnish junior high school. The participants (N = 53) were 13-15 years old. 30 of the participants were boys and 23 were girls. 16 of the participants played games daily and others were infrequent players. Mobile phone was the most popular gaming platform among the participants. However, most of the participants used to also play games on computers and consoles. Tablets and hand consoles were used quite rarely to play games.
The game playing session was organized during a regular school day. First the participants were introduced to the Labyrinth game in small groups (3-5 players). In practice, participants were shown how the game is played and the idea of the game was presented. Participants played the game for approximately 10-15 minutes depending on how well the game proceeded. The playing behavior was observed and video-recorded. Finally, immediately after the playing session, participants filled in an online questionnaire (likert scale 1-6) about their playing experience. After that participants were shortly interviewed in groups.

5.1.3. Results of the study 1

The results showed (table 1) that the flow level experienced by the players was in medium level (M = 3.57, SD = .81) and not as high compared to when playing educational games [see e.g. 7, 39, 40-42]. From table 1, we can see that playability (M = 2.77, SD = 1.67), rewarding experience (M = 2.89, SD = 1.17), and feedback (M = 2.93, SD = 1.21) dimensions scored the lowest. On the other hand loss of self-consciousness (M = 5.07, SD = .95), clear goals (M = 4.09, SD = 1.08), and concentration (M = 3.93, SD = 1.13) scored the highest.

Table 1. Means and standard deviations of flow dimensions (N = 53)

Flow dimension	M	SD
Challenge – skill balance	3.42	1.13
Clear goals	4.09	1.08
Feedback	2.92	1.21
Playability	2.77	1.17
Sense of control	3.63	1.34
Rewarding experience	2.89	1.17
Concentration	3.93	1.13
Loss of self-consciousness	5.06	.95
Time distortion	3.41	1.52
Flow experience (construct)	3.57	.81

It was surprising that the loss of self-consciousness dimension scored the highest. In many games, especially exergames when played within a public setting, players tend to feel shy and conscious of the others watching their performance. In this study however, players lost their self-consciousness. In fact, it seemed that the major goal of the players was not to solve the puzzle, but to learn together how to control the game characters. So, in the way, everybody had problems and they encouraged each other to test different styles to control the game. This has probably positively affected concentration and goal dimensions, although players focus were different than originally designed. However, we assume that the problems of the user interface reflected negatively by the other flow dimensions.
Observations conducted via video analysis and discussions with players revealed that players had problems in figuring out how to control their game character. Although the players were informed about how to control the game, they still encountered problems. In fact one of playability's big dilemmas is the common gap between designers' and players' mental models. Mental model (or conceptual model) is a player's explanation of how a game works. It is a simplified version of the designer's model that involves all the facts of the game. In our case, the players' mental model about the Labyrinth Run game largely differed from the designer's model reflected by the game, which led to playability problems. The discussions with players revealed that the earlier playing experiences with camera-based exergames, such as Microsoft Kinect games caused complication to the formation of the right kind of mental model concerning the Labyrinth game. Almost half of the players did not understand that their posture and moving direction had no effect on the behavior of their game

character. Players tended to run for example to the right and wondered why their character headed left in the game. This means that players assumed that their movements are mapped directly to the game world, which was not the case; 1) game character 's walking speed feeds of a player's motion intensity and the character changes direction only after collisions with walls or obstacles; 2) Game character jumps to the direction that it is facing when a player jumps, but not to direction that a player actually jumps. In earlier studies [43], we have successfully used similar game mechanics, but the controlling was a bit different. For example, in the Diamond Hunter game, the characters moved all the time and the players would only control the jumping. In that game, users did not face any playability problems. Despite the bad playability, the sense of control was quite high, which indicates that players believed that they could learn to control the game in the long run.
Furthermore, the results revealed that frequent players' did not experience high flow (M = 3.44, SD = 1.17) as infrequent players (M = 3.63, SD = .60). However, the difference is not statistically significant t(51) = .75, p = .46. When considering isolated flow dimensions, the biggest differences between frequent players and infrequent players are in feedback t(51)=2.84, p=.006, rewarding experience t(51)=1.64, p=.11 and playability t(51)=1.53, p=.13 dimensions. The gender did not affect either the flow construct t(51)=.69, p=.49 or isolated flow dimensions. Very small differences were found in playability, feedback and rewarding dimensions.

5.2. Study 2: TeamBoost

The new version of the Labyrinth Run game called TeamBoost was developed according to the findings of study 1. The aim was to offer more content with more diversity, increase cooperative elements and tackle the most obvious playability issues of Labyrinth Run. The educational objectives of the TeamBoost are the same as in Labyrinth Run dealing with soft skills. Furthermore, the design patterns for cooperative games [32] were more exhaustively applied and guidelines for movement-based games also informed the design solutions [34].

5.2.1. Description of TeamBoost game

TeamBoost has a stronger story theme compared to Labyrinth Run as players control a group of secret agents in a laboratory, mansion or factory settings (Fig. 3.). In order to better match the challenge of the level to the skills of the players, the levels are freely selectable and they have a difficulty rating (0-3 stars). Difficulty ranges from a tutorial level, where some basic tasks are performed with onscreen instructions, to really challenging levels, which require some serious thinking and problem solving even for experienced players. A couple of new winning conditions were also added.

Figure 3. A screenshot of the TeamBoost game (zoomed in); one player has built a bridge to help another player to overstep the gap.

TeamBoost uses the same *shared goal* design pattern [32] as Labyrinth Run. The aim is to complete the given mission as a team. Removing the competition aspect in intragroup level emphasized the meaning of a shared goal. In practice, both personal and group scores were discarded. Furthermore, we developed missions that can be completed only if every agent exits the building. In these missions, players have to collaboratively decide when each agent can exit the building. Because this action is permanent, bad decisions can lead to situations that the other agents are unable to exit if the skills or the presence of the agent, who left the building would have still been needed.
TeamBoost offers same cooperation mechanics as Labyrinth Run with the same interactive items, such as treadmills and scale platforms. Furthermore, some new items including elevators and power switches are included to add variation to puzzles. In addition to the new items, '*synergies between goals'* design pattern [32] was given more thought through the level design. Although Labyrinth Run had the required mechanics implemented, its levels had only a few tasks that really relied on cooperation. The levels of TeamBoost contain more tasks that require multiple agents to complete. For instance, two agents have to operate doors in correct order to allow a third agent to pass through a platform or every agent has to stand on their separate weight triggers to open the exit door.
The use of complimentary design pattern [32] was improved most in TeamBoost. In Labyrinth Run, the character attributes were identical apart from the weight difference (which is not used in TeamBoost) and players did not notice the difference in the characters' behaviour. In TeamBoost the characters are all different. One of the characters is a basic agent with no extra functionality but each of the other characters has a different special skill such as bridge building. This elevates the opportunities for cooperative game play to a higher level. All the special skills are designed for assisting other characters reflecting the '*abilities that can only be used on another player'* design pattern [32]. For example, ladder agent can set up a ladder for other agents to climb but cannot climb these ladders himself. The '*synergies between abilities*' design pattern [32] is used occasionally. Sometimes special skills are required to be used in synergy, for example a rope agent stands on a bridge, built by a bridge agent, and uses his rope to lift another agent up to the bridge. Cooperation is also facilitated by character limitations. For example, the bridge agent cannot jump and is often dependent on other agents to move from a platform to another.
The controlling method of Labyrinth Run, relying only on physical movements, proved to be too inaccurate, too intense, and confusing for players reflected especially by low playability and feedback scores. Players had difficulties to time their jumps and the turning mechanic was slow and confusing. Thus in TeamBoost, physical movement is restricted to running still to move the agent and to operate some interactive items such as treadmills and ladders. The rest of the actions are performed through buttons on a mobile device. Turning and jumping, which caused the most confusion amongst players in Labyrinth Run, are now operated with buttons. Likewise there is a button for activation of character's special skill (if available) and a button for interacting with items on the level. The aim of implemented user interface solutions is to make the controlling of the game easier to enable players to concentrate on the main tasks of the game. In other words, increasing the playability and sense of control dimensions as well as direct the perceived challenge to puzzles and collaboration instead of game controls.

5.2.2. Participants and procedure

The study was conducted in spring 2014 at the same Finnish junior high school. The participants (N = 49) were 13-15 years old. 14 of the participants were girls and 35 were boys. 23 of the participants played games daily and others were infrequent players. Mobile phone was the most popular gaming platform among the participants. However, most of the participants used to also play games with computers, consoles and tablets. Hand consoles were used quite rarely to play games.
The game playing session was organized during a regular school day. There were two parallel gaming areas in the same classroom and thus two games could be played simultaneously. Participants formed groups of four people and the groups were introduced to the TeamBoost game. In practice, the idea of the game was explained and participants were shown how the game is played.
Participants played the game for approximately 15-25 minutes depending on how well they proceeded. First, they had a practice session with an introductory level. After that, teams could select the level that they wanted to play. Two researchers observed the game sessions. Each team played approximately two levels. However, some of the teams played even three levels. Finally, immediately after the playing session, participants filled in an online questionnaire about their playing experience and they were informally interviewed.

5.2.3. Results of the study 2

The results show that the flow level experienced by the players were quite high (M = 4.18, SD = .99) and correspond to flow scores experienced in several other educational games [e.g. 25, 39-42]. From table 2, we can see that challenge – skill balance (M = 4.64, SD = 1.16), sense of control (M = 4.47, SD = 1.23), clear goals (M = 4.37, SD = 1.30), and loss of self-consciousness (M = 4.36, SD = 1.56) dimensions scored the highest. Similar to study 1, feedback (M = 3.64, SD = 1.21) and playability (M = 3.80, SD = 1.19) dimensions scored the lowest. For example, the use of ladders was unintuitive and the game did not provide clear feedback about activities performed and possibilities to act. Nevertheless, the scores of feedback and playability dimensions improved significantly from the first iteration. The gender did not affect either the flow construct t(51)=.06, p=.92 or isolated flow dimensions.

Table 2. Means and standard deviations of flow dimensions and flow construct (N = 49)

Flow dimension	M	SD
Challenge – skill balance	4.64	1.16
Clear goals	4.37	1.30
Feedback	3.64	1.21
Playability	3.80	1.19
Sense of control	4.47	1.23
Rewarding experience	4.00	1.45
Concentration	4.23	1.35
Loss of self-consciousness	4.36	1.56
Time distortion	4.08	1.21
Flow experience (construct)	4.18	.99

The observation notes indicated that players learned how to control TeamBoost much faster than Labyrinth Run game. Players also understood the meaning of cooperative elements and could use their agents' special skills in reasonable ways. Overall, the shared goals were clearer. Hence, the players could clearly concentrate on solving the puzzles in teams instead of rushing around in the game world without strategies. Players could avoid overloading their cognitive or physical capacity by sequencing the cognitive and physical gameplay. For example, momentarily players stopped all physical activities to communicate with their team and conceptually solve the puzzle and then performed the needed physical activities to test their solution. Such planning and negotiation phases did not appear in study 1. We also noticed that the teams that had a clear team leader managed to solve the levels most effectively.

Furthermore, the results indicated that frequent players' experienced little higher flow (M = 4.34, SD = 1.07) than infrequent players (M = 4.03, SD = .91). However, the difference is not statistically significant t(47)=-1.07, p=.29. In spite of that, this finding is controversial when compared with the findings of the study 1 in which infrequent players experienced higher level of flow. The biggest differences between infrequent players and frequent players were in goal t(47)=-1.35, p=.185, loss of self-consciousness t(47)=-1.157, p=.25 and feedback t(47)=-1.12, p=.27 dimensions. The interviews indicated that frequent players liked the game and they saw a lot of potential in TeamBoost if the small playability problems are fixed and more informative feedback is provided for players. In general, these results indicate that the quality of TeamBoost was much better than Labyrinth Run. In the next section, results of both studies are summarized and the game features that will be implemented to improve TeamBoost are presented.

5.3. Summary of findings and future development

The results clearly demonstrated that players appreciated the TeamBoost game more than Labyrinth Run game. This concludes that the modifications that were made to the first prototype were successful. Figure 4 illustrates the increment of flow scores dimension by dimension. The overall improvement was statistically significant t(51)=3.39, p=.001. The most significantly improved the challenge t(51)=5.41, p=.000, the playability, t(51)=4.38, p=.000, and the rewarding experience t(51)=4.24, p=.000 dimensions. Only the loss of self-consciousness dimension scored lower in TeamBoost condition. We assume that the reason for this was the different abilities of each game character that could be used to aid teammates. Some of the players felt that this feature increased the individual pressure and they started to think about how their teammates would evaluate their playing behaviour. In spite of this, the use of complementarity design pattern is important, because it facilitates the cooperation between players. In fact, the goal of the game is to create some pressure, teach players to work in teams and handle their feelings when experiencing pressure.

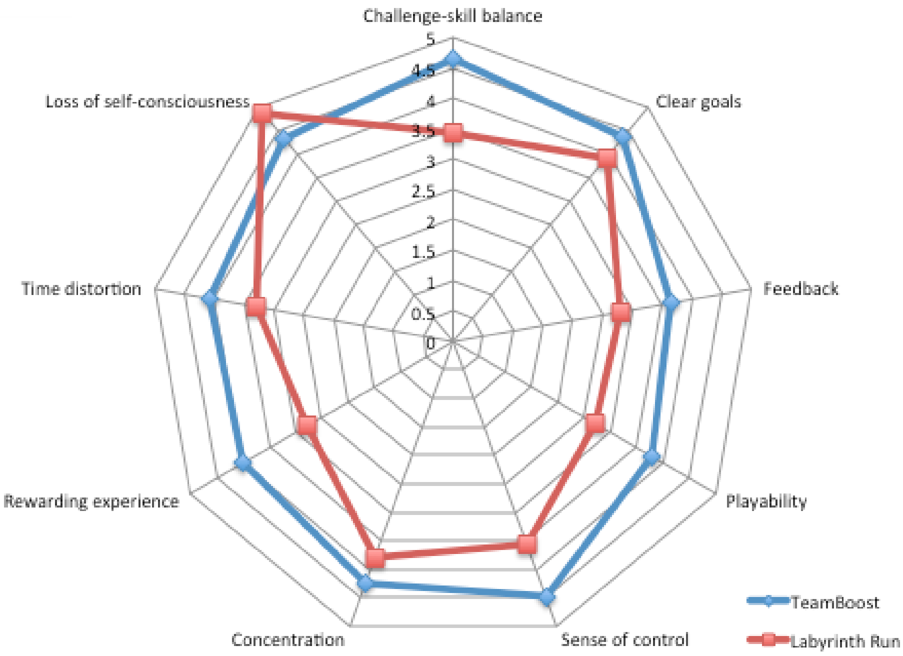

Figure 4. The means of flow dimensions in TeamBoost and Labyrinth Run games

Although the players liked TeamBoost a lot, several improvements are required based on the gathered results. For instance, the characters will be controlled horizontally by tilting the mobile device instead of running still. This will make the controlling of the game easier and players can concentrate on solving the puzzles. This will of course decrease physical activities. Thus, more interactive items that require physical movements have to be implemented in order to retain the exergame aspect. In the next version, the following elements will require physical movements: ladders, rope, treadmill, movable boxes, and elevator. All in all, the decrement of physical activities is reasonable because the intended fatigue should be used only if it is an important challenge of the game, which is not the case in TeamBoost [34]. The overall idea is to add physical activities to elements that do not require accuracy in order to facilitate playability of the game.

As identified by the players, the current feedback provision is not constructive and players did not always know what they can do in the game and how they progress against goals. For example, how long a bridge can a player build, how long a rope should be used or can ladders be created for a Rope agent. Currently, the feedback on player's actions is mainly displayed through animations, such as an agent spreading his arms when a special skill could not be used. However, the feedback does not provide any reason to why the skill cannot be used. Thus, we will focus on refining the feedback mechanism by providing more cognitive and unambiguous feedback. For example, a transparent bridge of a maximum length will be shown when a player is attempted to build a bridge over a gap too wide. Also the message box of the mobile application that is used to control the agent will be utilized as a feedback channel. Currently, it mainly displays the status information about network connections. As an addition, it will be used to provide for example hints about the next tasks on the level. Feedback had clear connection with several flow dimensions and thus the improved feedback mechanisms should have a positive influence also on other flow dimensions, particularly on playability ($r = .51$), sense of control ($r = .43$) and clear goals dimensions ($r = .48$).

Furthermore, general content creation will be continued. New character types and more interactive items will be implemented to add variety to levels and increase the opportunities for cooperation combinations. Also the role of the basic character that does not have any special skill will be rethought, because on some occasions the player controlling the basic agent felt disappointed that he or she could not contribute more in the game.

6. Conclusions

In this paper, we considered the usefulness of flow experience in evaluating the quality of serious games (exergames). We reported the results of a game design process in which two prototypes (iterations) of a collaborative exergame were developed and studied. In general, the results indicated that the measurement of the flow experience can reveal weaknesses of the game and consequently aid the design process. For example, the first prototype studied, Labyrinth Run, did not create high flow experience in players. Especially, the playability of the game was low and the game did not provide adequate feedback to players. The measurement of flow provided us a grounding for the redesigning of the game. The results of the second study (second iteration) showed that TeamBoost was appreciated much more in terms of flow compared to Labyrinth Run. This indicates that the modifications that were made to Labyrinth Run informed by the flow measurements were successful. However, the flow scores of TeamBoost indicate that there are still issues in the game that may disrupt playing experience. Thus, we presented several improvements that will be implemented on TeamBoost game. Further development work will focus especially on creating more intuitive and fluent game controller that frees cognitive resources to problem solving processes, balancing both physical and cognitive workloads and implementing a more constructive feedback system.

To summarize, based on these findings we argue that flow experience can be used to evaluate the overall quality of a playing experience and it provides a structured approach to consider the quality of a game. However, it does not provide detailed information about the weaknesses or highlights of the game and thus complementary methods is needed to identify these issues, which will further inform design improvements. In this research, we conducted interviews and observations as complementary sources and they provided deeper and more constructive information about the quality of the studied games. The results also indicated that flow experience was independent of gender, which supports its use in quality measurement.

This research has also several limitations. Firstly, the participants played the game for only a limited time (15-25 minutes), which may have influence on the experienced flow level. For example, Lai, Wang and Yang [44] found that more frequent play of exergames and longer use of the exergame consoles may increase the level of flow and enjoyment. On the other hand, in the era of mobile games, children are used to play games that they can understand in seconds. Even a couple of minutes of a learning curve may be too long and even a good and an effective educational game can fail in the market. Thus, the first impression and the first playing session can reveal a lot about the potential of a game and the short playing time of the study is justifiable. Secondly, the study concentrated solely on measuring flow experience, although for example the measurement of team commitment would have provided interesting information about the relation between collaborative gameplay and flow. Thus, in the future we aim to study social dimension more deeply and also evaluate the usefulness of TeamBoost in teaching soft skills.

Acknowledgements

This work has been co-funded by the EU under the FP7, in the Games and Learning Alliance (GALA) Network of Excellence, Grant Agreement nr. 258169 and by Academy of Finland under Research Programme on the Future of Learning, Knowledge and Skills (TULOS nr. 274240).

References

[1] Procci, K., Singer, A. R., Levy, K. R., Bowers, C., "Measuring the flow experience of gamers: An evaluation of the DFS-2". Computers in Human Behavior, *28*(6), 2306-2312, 2012.

[2] Wirth, W., Hartmann, T,. Bocking, S., Vorderer, P., Klimmt, C., Schramm., H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., Biocca, F., Sacau, A., Jäncke, L., Baumgartner, T., Jäncke, P., "A process model of the formation of spatial presence experiences". Media Psychology, 9, pp. 493-525, 2007.

[3] Kiili, K., de Freitas, S., Arnab, S. Lainema, T., "The Design Principles for Flow Experience in Educational Games". Procedia Computer Science, 15, 78-91, 2012.

[4] Cowley, A. D., Minnaar, G., "New generation computer games: watch out for Wii shoulder", BMJ: British Medical Journal, 336(7636), 110, 2008.

[5] Sweetser, P., Wyeth, P., "GameFlow: a model for evaluating player enjoyment in games". Computers in Entertainment, 3(3), 1-24, 2005.
[6] Sinclair, J., Hingston, P., & Masek, M., "Exergame development using the dual flow model". In Proceedings of the Sixth Australasian Conference on Interactive Entertainment (11), ACM, 2009.
[7] Bressler, D. M., & Bodzin, A. M., "A mixed methods assessment of students' flow experiences during a mobile augmented reality science game". Journal of Computer Assisted Learning, 29(6), 505-517, 2013.
[8] Barzilai, S., & Blau, I., "Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences". Computers & Education, 70, 65-79, 2014.
[9] Liu, T. Y., "Using educational games and simulation software in a computer science course: learning achievements and student flow experiences". Interactive Learning Environments, 1-21, 2014 (ahead-of-print).
[10] Csikszentmihalyi, M., "Flow: The Psychology of Optimal Experience". Harper Perennial, New York, 1991.
[11] Hsieh, Y. H., Lin, Y. C., & Hou, H. T., "Exploring the role of flow experience, learning performance and potential behavior clusters in elementary students' game-based learning". Interactive Learning Environments, 1-16, 2013 (ahead-of-print).
[12] Ma, Y., Williams, D., & Prejean, L., "Designing an Electronic Educational Game to Facilitate Immersion and Flow". Journal of Interactive Learning Research, 25(1), 27-49, 2014.
[13] Hou, H. T., & Li, M. C., "Evaluating multiple aspects of a digital educational problem-solving-based adventure game". Computers in Human Behavior, 30, 29-38, 2014.
[14] Ketamo, H., Kiili, K., "Conceptual change takes time: Game based learning cannot be only supplementary amusement". Journal of Educational Multimedia and Hypermedia, 19(4), 399-419, 2010.
[15] Miller, G., "The magical number seven, plus or minus two: some limits on our capacity for processing information". Psychological Review, 63, 81-97, 1956.
[16] Kiili, K., Perttula, A., "A design framework for educational exergames", In Freitas S., Ott, M, Popescu, M.M. and Stanescu, I. (Eds.), New Pedagogical Approaches in Game Enhanced Learning: Curriculum Integration, US: IGI Global, 2013.
[17] Mueller, F., Edge, D., Vetere, F., Gibbs, M. R., Agamanolis, S., Bongers, B., & Sheridan, J. G., "Designing sports: A framework for exertion games". In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Vancouver, Canada: ACM Press, 2011.
[18] Connolly, C. T., & Tenenbaum, G., "Exertion–attention–flow linkage under different workloads". Journal of Applied Social Psychology, 40(5), 1123–1145, 2010.
[19] Fogg, B. J., "Persuasive technology: Using computers to change what we think and do". San Francisco, CA: Morgan Kaufmann, 2002.
[20] Sweller, J., Van Merriënboer, J., & Paas, F., "Cognitive architecture and instructional design". Educational Psychology Review, 10(3), 251–296, 1998.
[21] Tenenbaum G., "A social-cognitive perspective of perceived exertion and exertion tolerance". In: Singer, R.N.,Hausenblas, H. and Janelle, C. (Eds.), Handbook of sport psychology, New York: Wiley, 2001.
[22] Tenenbaumb, G., Connollya, T.C., "Attention allocation under varied workload and effort perception in rowers". Psychology of Sport and Exercise, 9(5), 704-717, 2008.
[23] Scott, L. M., Scott, D., Bedic, S. P., Dowd, J., "The effect of associative and dissociative strategies on rowing ergometer performance". The Sport Psychologist, 13, 57-68, 1999.
[24] Zagal, J. P., Rick, J., & Hsi, I., "Collaborative games: Lessons learned from board games". Simulation & Gaming, 37(1), 24-40, 2006.
[25] Azadegan, A., & Harteveld, C., "Work for or Against Players: On the Use of Collaboration Engineering for Collaborative Games". In proceedings of the 9th International Conference on the Foundations of Digital Games, 2014.
[26] Kiili, K, Tuomi, P. & Perttula, A., "Exerbraining for schools: Combining body and brain Training.". Procedia Computer Science, 15, 163-173, 2012.
[27] Sedano, C. I., Carvalho, M. B., Secco, N., and Longstreet, C. S., "Collaborative and cooperative games: Facts and assumptions". In Proceedings of the International Conference on Collaboration Technologies and Systems (370-376), 2013.
[28] Malone, T., Lepper, M., "Making learning fun: a taxonomy of intrinsic motivation for learning". In Snow R. E. and Farr, M. J. (Eds.), Aptitude, Learning, and Instruction, Hillsdale, N. J: Lawrence Earlbaum, 1987.

[29] Bruffee, K., “Sharing our toys- Cooperative learning versus collaborative learning”. Change, 12-18, 1995.
[30] Romero, M., Usart, M., Ott, M., Earp, J., de Freitas, S., Arnab, S., “Learning through playing For or against each other?”. Promoting Collaborative Learning in Digital Game Based Learning, ECIS 2012.
[31] Moisant, P., “Course sans Gagnant”. Retrieved 05-08, 2014, from http://www.patrickmoisan.net/copains/course_sans_gagnant.html, 2005.
[32] Rocha, J. B., Mascarenhas, S., Prada, R., “Game Mechanics for Cooperative Games”. In ZDN Digital Game, 2008.
[33] March, S.T., Smith, G.F., “Design and natural science research on information technology”. *Decision Support Systems*, 15, 251-266, 1995.
[34] Mueller, F., & Isbister, K., “Movement-based game guidelines”. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems (2191-2200), ACM, 2014.
[35] Weber, R., Tamborini, R., Westcott-Baker, A. Kantor B., “Theorizing flow and media enjoyment as cognitive synchronization of attentional and reward networks”. Communication Theory, 19, 397-422, 2009.
[36] Kiili, K., Lainema, T., ”Foundation for Measuring Engagement in Educational Games”. Journal of Interactive Learning Research, 19(3), 469-488, 2008.
[37] Jackson, S. A., Eklund, R. C., “Assessing flow in physical activity: The Flow State Scale-2 and Dispositional Flow Scale-2”. Journal of Sport & Exercise Psychology, *24*(2), 2002.
[38] Pasch, M., Bianchi-Berthouze, N., van Dijk, B., & Nijholt, A., “Movement-based sports video games: Investigating motivation and gaming experience”. Entertainment Computing, 1(2), 49-61, 2009.
[39] Kiili, K., Lainema, T., “Power and Flow Experience in Time-Intensive Business Simulation Game”. Journal of Educational Multimedia and Hypermedia, 19(1), 39-57, 2010.
[40] Brom, C., Bromová, E., Děchtěrenko, F., Buchtová, M., & Pergel, M., “Personalized messages in a brewery educational simulation: Is the personalization principle less robust than previously thought?”. Computers & Education, 72, 339-366, 2014.
[41] Fu, F. L., Su, R. C., & Yu, S. C., “EGameFlow: A scale to measure learners’ enjoyment of e-learning games”. Computers & Education, 52(1), 101-112, 2009.
[42] Hsiao, H. S., Chang, C. S., Lin, C. Y., & Hu, P. M., “Development of children's creativity and manual skills within digital game-based learning environment”. Journal of Computer Assisted Learning, 2014.
[43] Kiili, K., Perttula, A., Tuomi, P., “Development of multiplayer exergames for physical education”. IADIS International Journal on WWW/Internet, 8(1), 52-69, 2010.
[44] Lai, Y. C., Wang, S. T., & Yang, J. C., “An investigation of the exergames experience with flow state, enjoyment, and physical fitness”, In proceedings of the Conference on Advanced Learning Technologies (ICALT) (58-60). IEEE, 2012.