

Comparing Voluntary and Mandatory Gameplay

Esther Kuindersma¹, Jelke van der Pal¹, Jaap van den Herik² and Aske Plaat²

¹Netherlands Aerospace Centre (NLR), Amsterdam, the Netherlands
Esther.Kuindersma@NLR.nl

²Leiden University, Leiden Institute of Advanced Computer Science,
Leiden, the Netherlands

Abstract

Gameplay is commonly considered to be a voluntary activity. Game designers generally believe that voluntary gameplay is essentially different from mandatory gameplay. Such a belief may be a challenge for serious games, as instruction is usually mandatory. The article describes the outcomes of two experiments on the impact of voluntariness on the learning effect and enjoyment of a serious game. In the first experiment freedom of choosing to play a serious game was studied, with participants who had volunteered to participate. The results suggested that, contrary to the opinion of many game designers, being required to play a serious game does not automatically take the fun out of the game. The second experiment had voluntary participants and mandatory participants, who had to participate as part of a homework assignment. The outcomes show that mandatory participants enjoyed the game as much as the voluntary participants, even if they had to play the game for a minimum required time. These studies indicate that mandatory gameplay does not reduce enjoyment and learning effect.

Keywords: serious games; learning effect; enjoyment; mandatory play

1. Introduction

Serious games are “games that do not have entertainment, enjoyment or fun as their primary purpose” [1]. Since the turn of the century, the research of this type of game has developed into a topic of considerable substance in educational research, although studies mostly report only slightly positive outcomes [2]. Specifically, the effectiveness of game-based learning is a topic of interest, and many studies focus on the comparison of the effects of serious gaming to those of traditional learning methods [3]. However, traditional learning methods are generally mandatory by nature, whereas gaming is assumed to occur in a voluntary setting: the player has a choice to play or not to play. The literature reports that freedom of choice has a positive effect on motivation and participation [4, 5]. As these last factors are also thought to have an influence on learning effect, it is plausible that voluntariness will have a positive effect on learning outcomes too.

In our literature search no serious gaming studies were found that dealt with the impact of freedom of choice on motivation and learning outcomes (see section 2). The study further serves to fill this hiatus by determining whether, and to what extent, gameplay and learning effect of a serious game are affected by the player’s freedom to choose to play the game. The article is organized as follows: section 3 focuses on the first experiment using an informal, non-educational setting, in which voluntary participants were randomly assigned to either play the serious game for at least a minimum amount of time or to have the freedom to choose how long to play the game. The first experiment has been repeated with the addition of mandatory participation conditions to create a stronger mandatory condition of participation. This second experiment, described in section 4, was set in a formal educational environment and had a larger sample size to compare voluntary participants with participants with an obligation toward their teacher or school. The outcomes of both experiments combined will be discussed and compared in section 5, followed by our conclusion (section 6).

2. Background of the Study

The use of games in training is anything but new. This practice has been around for centuries [6]. While it had been used in different contexts before [7], the term ‘serious game’ was first used in relation to instruction in 1970 by Abt [8]. In his view, the instructional aspect of games does not have to be incorporated into the game itself, but can also be part of the context. In 2002 the meaning of ‘serious games’ moved toward digital games [9]. Nowadays, serious games are defined as (digital) games with a main purpose other than entertainment, enjoyment or fun [1]. Serious games with an educational main purpose are also referred to as instructional games or as the instruments of game based learning.

Most definitions of serious games originate directly from game definitions in general. Games are defined by Salen and Zimmerman [10] as “systems in which players engage in an artificial conflict, defined by rules that result in a quantifiable outcome” and by McGonigal [11] as “activities with a goal, rules, a feedback system, and voluntary participation”. The words “voluntary” and “free” are common in definitions of games that are used by scholars on gameplay and games [1], [12, 13, 14, 15]. Although there is not a universally used definition, used by all game designers, there is general agreement on the main characteristics of games as well as the understanding that a game does not necessarily need to satisfy all of these. In most definitions games are thought to have rules, goals, a storyline and outcomes, and they offer interaction, feedback and competition. Furthermore, and critically important: games are played voluntarily and they are fun, or as they can be frustrating at times – at least they are ‘immersive’ or ‘engaging’.

The term “serious games” is sometimes considered an oxymoron, mostly on account of the fun characteristic of games. Games are fun by definition and therefore cannot be serious at the same time [16]. In contrast, Huizinga [12] stated that play is a serious activity, but that fun and serious do not necessarily exclude one another. According to Huizinga, games are non-productive and separate from the real world, whereas serious games have specific learning objectives related to life or work skills [14]. Caillois [13] has stated that it ceases to be play when this play of a game is forced and thus, games should be played voluntarily. Yet serious games are often meant to be instructional and instruction is typically non-voluntary [13]. These paradoxes may reflect on a player’s attitude, which in turn may have an impact on the actual learning effect of the serious game. Giving students a choice in whether to play the serious game or not, may result in a more positive attitude, a higher engagement and more time spent in the game [17].

Offering learners a choice in their assignments empowers them to take control, providing ownership of their learning process and motivating them to be engaged. This leads to increased interest and increased time spent on the chosen assignment [18]. The freedom to choose what, when, and how to contribute in the learning process can motivate learners to actively participate and accomplish more [19].

In a study of forced play, Heeter et al. [4] found that non-gamers, with little or no experience with digital games, are likely to be at a disadvantage in serious gaming, as obtaining the intended effect of a serious game depends on how well the game is played. The negative affect that non-gamers experience in a game is expected to interfere with learning or with the cognitive benefits. Their study also showed that resistant players have less attention for the game they have to play, and that they experience less positive and more negative feelings about that game. They would not play the game if they did not have to. Heeter et al. [4] concluded that serious games are least effective for players who dislike a game, and most effective for those who like it. Gauthier et al. [20] found no effect of gaming habits on voluntary use of a serious game in an unstructured, informal learning environment. They did find a positive effect of studying habits. Students reporting good studying habits using the game as a study tool, completed more tasks than students with good studying habits in the control group using a non-game tool.

Closely related to freedom of choice is the topic of consent. Mollick and Rothbard [5] examined the role of consent as a psychological response to “mandatory fun” in gamification in the work environment. They found that consent to a game significantly increased the employees’ positive affect, while resistance resulted in a decrease in positive affect and performance. They identified two sources of consent: (1) employees who play games outside of work are more likely to consent to them in other settings, and (2) individuals who are allowed to choose which game to play, show higher levels of consent and perceived control. This may coincide with the freedom to choose to play a serious game or not. From these two findings we are inclined to believe that playing a serious game voluntarily will increase positive affect and possibly performance.

Based on the motivating aspect of choice and the original definition of games we formulate the leading research question of the study: “Does freedom of choice (voluntariness) have a positive impact on the learning effect of serious games?”

To answer this question we conducted two experiments to determine whether, and to what extent voluntariness influences the gameplay and the learning effect of a serious game. In the article we describe these experiments.¹

3. First experiment: Voluntary versus Mandatory Gameplay in an Informal Setting

The first experiment was set up to determine whether voluntarily using a game as a learning tool will result in higher performance. The experiment consisted of (1) a pre-experiment questionnaire, (2) a selection of learning materials and a test, and (3) a post-experiment questionnaire. The design and the procedure will be described in sections 3.1 and 3.2, followed by the participants in section 3.3. Section 3.4 focuses on the materials that were used in the experiment. The results are presented in section 3.5 and discussed in section 3.6.

3.1. Design

The study used an experimental design with two independent conditions, to which participants were randomly assigned: (1) a Voluntary Gaming (VG) group that was free to choose how long to play the game or not to play the game at all, and (2) a Mandatory Gaming (MG) group that had to actively play the serious game for a minimum of 10 minutes.

The dependent variables were game score, test score and time spent playing the game. All participants were Voluntary Participants (VP) to the experiment. Participants were offered a chance to win a € 100 gift certificate. Chances of winning were related to completing all parts of the experiment, not to personal results.

3.2. Procedure

The experiment’s procedure and materials are shown in Fig. 1. Participants could complete all parts at their own convenience through a webpage with the experiment instructions and materials. In an online questionnaire, participants were asked about demographic information and their personal motivation for participating.

The participants were then asked to study the text-based materials and play the game, if applicable for their respective gameplay condition. They did so at their own pace and were free to proceed to the test when ready.

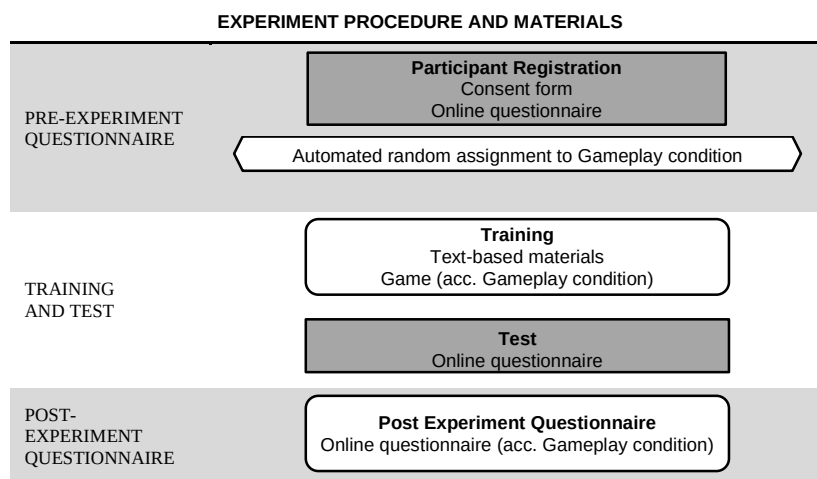


Figure 1. Schematic representation of the procedure and corresponding materials

¹ The first experiment has been reported previously [21].

3.3. Participants

Participants were recruited through social media and by personal invitation. A total of 64 persons registered for the experiment and completed an online questionnaire with demographic information and levels of motivation and prior knowledge. They were randomly assigned to one of two groups, resulting in a Voluntary Gameplay (VG) group of 29 participants and a Mandatory Gameplay (MG) group of 35 participants (Table 1). Nineteen participants completed the training and the test. Enquiries after the experiment ended, show that many participants discontinued their participation due to other priorities; such as work or social obligations. There were nine completed responses from voluntary players and ten from mandatory players. The groups did not differ significantly in terms of sex, age and interest in gaming.

Table 1. Distribution of participants over Voluntary and Mandatory gameplay conditions

Voluntary Participation	Gameplay		Total
	Voluntary (VG)	Mandatory (MG)	
Registered	29	35	64
Completed	10	9	19

In the pre-experiment questionnaire, participants indicated their prior knowledge on a scale of ten. This resulted in a mean score of 4.2 (SD = 2.4) with no significant difference between groups. However, in the pre-experiment question pertaining to their motivation to participate, mandatory players indicated to be more motivated than voluntary players (One-way ANOVA: $F(1,17) = 9.3$, $p < .05$).

3.4. Materials

For the first experiment a set of materials was developed, consisting of (1) a pre-experiment questionnaire with a short assessment of prior knowledge, (2) written materials, a serious game and a test about cloud identification and risk assessment, and (3) a post-experiment questionnaire. We will first discuss the game, followed by the written materials, the test and the questionnaires.

3.4.1. CloudAtlas Game

The CloudAtlas game itself is identical for both the voluntary (VG) and mandatory (MG) gameplay conditions, but mandatory players (MG) had to play at least 10 minutes for the test to become available. The game is played in an internet browser using the keyboard as the input device. The objective is to fly an aircraft as far as possible (Fig. 2). During flight the player encounters the ten types of clouds that have been addressed in the written materials. Applying their knowledge about clouds and possible hazards, the players must make decisions on how to respond. They can (1) fly through a cloud, (2) go over or under it, or (3) land the aircraft to wait for the danger to pass. Clouds can lead to three hazards to the aircraft: icing, turbulence and lightning. These hazards are visualized on the screen and have an impact on the game by increasing the aircraft's fuel consumption. Some hazards may be avoided by flying at high altitude, which requires oxygen. The player starts with a limited supply of oxygen. During the game extra fuel and oxygen can be picked up to extend the flight. The player may also encounter balloons and flocks of birds. Collisions must be avoided, because they will end the game. In all other cases the game will end when the player runs out of fuel or oxygen. The distance traveled translates into a game score. Picking up boosters during flight adds to the score, while points are deducted for making unnecessary landings.

Each individual game is relatively short. The game is intended to challenge the players to improve their high score, thus providing repeated exposure to the cloud types and their consequences.

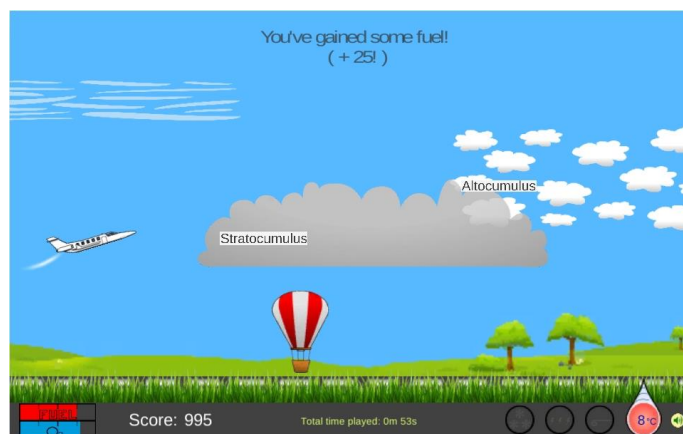


Figure 2. Screenshot of the game

3.4.2. *Written materials*

The written materials offer information about cloud classification, characteristics of different cloud types, possible hazards, and the impact of clouds on aviation. It shows drawings and photographs of different types of clouds. All participant groups had unlimited access to the same set of text-based materials.

3.4.3. *Test*

The test consists of eleven knowledge questions and seven application questions. The knowledge questions ask participants to reproduce cloud characteristics and recognize clouds from drawings and photographs. In the application questions, players apply their knowledge to a specific situation. For example, a picture of an aircraft and a particular type of cloud is presented with a number of possible routes (Fig. 3). Participants were asked to choose the best route, taking into consideration safety, comfort and efficiency. They were also asked to explain their reasons for choosing the specific answer. Questions were assigned different weights. Test scores were calculated as the percentage of points earned out of the maximum. In general, application questions were assigned higher weights than knowledge questions.

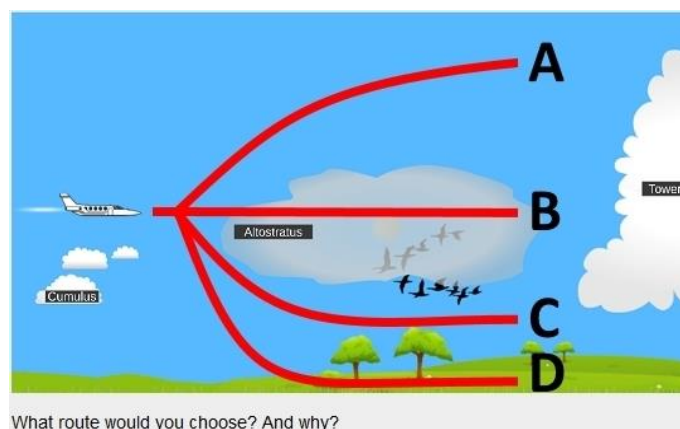


Figure 3. Test item: Application question

3.4.4. *Post-experiment questionnaire*

After the test, participants were presented with the post-experiment questionnaire. This questionnaire solicited more information about gaming preferences and personal motivation. The voluntary players were asked about the extent of the freedom of choice they experienced in choosing to play or not to play the game. The mandatory players were asked whether they would have played the game when given a choice.

The questionnaires have been constructed specifically for the study. No validated survey questions pertaining to voluntariness or game enjoyment had been found in literature. Several

questionnaire items used a ten point scale (1...10). Such a scale is easily understood across age groups and education levels [22], and provides appropriate data for analysis [23]. The use of an even scale avoids the neutral midpoint, forcing the participants to make a distinct choice for each item. Furthermore, the use of a ten point scale is common in both customer satisfaction surveys and game reviews.

3.5. Results

In the first experiment only 19 participants completed the experiment by taking the test; 16 of them played the game, three participants in the VG group chose not to play the game. Table 2 shows the means and standard deviations on game and test scores. One-way analysis of variance controlled for motivation (ANCOVA) was used to control for the possible effects of the group difference that was found on motivation, prior to the experiment. The analysis showed no significant differences for test scores between the groups.

Table 2. Means and SD of results for Voluntary and Mandatory gameplay conditions

Measure	Gameplay			
	VG n=10		MG n=9	
	M	SD	M	SD
Gameplay (min)	3.4	2.9	16.8	8.2
Test score (%)	44.9	11.3	48.7	18.3
Game score	1092	1085	2723	1332

We found a large variation in length of gameplay. Three VG participants chose not to play at all, while two MG participants played for more than half an hour.

Table 2 reveals that MG players played an average of 13.4 minutes longer than VG players ($F(1,17) = 23.50, p < .001$). The effect of the gameplay condition on the amount of time played using prior motivation as a covariate was significant, $F(1, 16) = 10.98, p < 0.01$.

Contrary to our expectations, we did not find a significant difference between the groups on test score and performance in the game.

We had anticipated seeing subsets of players in both the VG and the MG conditions: those who played only as long as required (up to 12 minutes) and those who continued playing (more than 12 minutes). Within the VG group, we also expected to find players who did not play at all, and players that only played to get an idea of the game by playing three tries or less.

In the mandatory group, 3 participants played less than 12 minutes, the other six played longer. Surprisingly, all ten players in the VG group played less than 12 minutes. Three players did not play at all and four players tried three times or less. Contrary to what we expected based on the literature, mandatory players played longer, not shorter than voluntary players.

In the post-experiment questionnaire all participants were asked how much they had enjoyed playing the game on a scale of ten ($M = 6.6, SD = 1.5$). Voluntary players did not enjoy the game significantly more than mandatory players.

We were surprised to find that voluntary players played shorter than mandatory players. As both groups reported equal enjoyment of the game, this difference does not need to be attributed to engagement. It seems that mandatory players do not feel much pressure and the obligation is merely experienced as a stimulus.

MG players were also asked how they felt about being obligated to play the game for a minimum amount of time. In general, participants were neutral about this ($M = 2.1, SD = .8$). When asked whether they would play the game if they were given a choice, almost 78% of the MG players indicated they would. This percentage is higher than the percentage of VG players who actually chose to play the game, which is 70%. Furthermore, the duration of intended gameplay indicated by the MG players was higher than the average time played by the VG players, which is 3.4 minutes. Although these differences were not significant (as a result of the small sample of participants), they are opposite to expectation and as such a remarkable result.

3.6. Discussion

By recruiting through social media we aimed to reach a large number of participants, but in fact the number of participants was small. The group difference on prior motivation would probably not have occurred with a larger sample size or a different assignment strategy (pair matching). From the 64

initial registrations, only 19 persons completed the experiment. This dropout rate may have influenced the results.

We recall that, contrary to what we expected based on the literature, mandatory players played longer, not shorter, and showed equal enjoyment. A possible explanation for the unexpected results may be that the level of obligation experienced by the MG group was minimal as a result of the voluntary participation in an on-line experiment. The relative difference between voluntary and mandatory gameplay in the experimental setting may manifest itself stronger in a formal setting, when lecturers provide gameplay in a voluntary or mandatory way. Mandatory participation to a regular course may provide for a more normal motivation setting for students, in which the effects of voluntary gameplay can be observed without self-selection issues. Therefore, a second experiment was conducted in order to challenge the outcome of the first experiment in a regular course setting.

4. Second experiment: Voluntary and Mandatory Gameplay in a Formal Setting

The second experiment served to determine whether voluntary gameplay has the same effects on learning outcome and enjoyment in a formal setting, where the primary motivation to participate is either voluntary or mandatory.

The experiment consisted of the same three parts as the first experiment (section 3.2). However, a new experimental design was used as a No Game control group was introduced. Also, in the second experiment higher education students were recruited to be either voluntary or mandatory participants to the experiment.

The new experimental design and the procedure will be discussed in sections 4.1 and 4.2. The participants are described in section 4.3. Then the adaptations in the materials from the first experiments are described in section 4.4, followed by the results (section 4.5) and discussion (section 4.6).

4.1. Design

The second experiment had an experimental design with six conditions (Table 3). First, participants were recruited either as Voluntary (VP) or Mandatory (MP) participants to the experiment. VP participants volunteered, while for MP participants the training was assigned as homework. Then, participants in both participation groups were randomly assigned to one of three gameplay groups: (1) a Voluntary Gaming group (VG) that was free to choose how long to play the game or not to play the game at all, (2) a Mandatory Gaming group (MG) that had to actively play the serious game for a minimum of 10 minutes, or (3) a No Game control group (NG) that had no access to the game.

Participation and gameplay formed the independent variables. The dependent variables were game score, test score and time spent playing the game.

As an incentive, VP participants were offered a chance to win a € 100 gift certificate. They were informed that chances of winning were related to completing all parts of the experiment, not to personal results. Participants receiving school credit for participating or doing it as a homework assignment were not eligible for the prize.

Table 3. Participation and Gameplay conditions

Participation	Gameplay		
	VG	MG	NG
VP	Volunteered for participation, free to choose to play	Volunteered for participation, minimum 10 minutes of gameplay	Volunteered for participation, no access to game
MP	Participation assigned as homework, free to choose to play	Participation assigned as homework, minimum 10 minutes of gameplay	Participation assigned as homework, no access to game

4.2. Procedure

The experiment's procedure and materials were similar to those of the first experiment (see Fig. 1). Participants could complete all parts online at their own convenience through a webpage with the experiment instructions and materials.

4.3. Participants

To create an obligation for students to participate, a formal learning setting was required in which lecturers assign the training as homework. For that reason participants were recruited through (applied) universities. Students were informed that the experiment related to aviation, but the focus on gaming remained undisclosed. Teachers and lecturers from sixteen faculties in twelve institutions assigned participation in the experiment as a homework task to their students, creating a sense of obligation from the students toward their teacher. If a teacher was unable to assign homework for any reason, (s)he informed their students about the experiment and invited them to participate without obligation.

Four teachers from four different institutions assigned homework to a total of approximately 90 students. One teacher posted the experiment on the ERAS network for Psychology students, who have to participate in experiments in order to get school credit. As these students are free to choose the experiments in which to participate, they are considered to be volunteers in the current study. Thirteen teachers from seven different institutions invited over 1000 students to participate voluntarily.

The experiment was completed by 93 participants; homework was assigned to 36 of them, 31 participated for school credit and 26 volunteered. Teachers did not report reasons for non-response. There were 42 men and 51 women with a mean age of 21.7 (SD = 3.6). The test groups did not differ significantly in terms of sex, age, motivation prior to the experiment and number of gamers. However, 19 participants spent less than 3 minutes studying the written materials combined with less than one minute playing the game, apart from the required time. They were not considered legitimate participants and their results were removed from the analyses.

Hence, 74 participants completed the experiment; 37 men and 37 women with a mean age of 21.9 (SD = 3.7). In total 46 students volunteered to participate, including the students who received school credit and 28 students participated as part of a homework assignment. This will be referred to as Voluntary Participation (VP) and Mandatory Participation (MP) respectively. All participants were randomly assigned to a gameplay condition, resulting in a Voluntary Gameplay (VG) group of 29 participants, a Mandatory Gameplay (MG) group of 21, and a control group without access to the game (NG) of 24. The distribution of participants is shown in Table 4.

Table 4. Distribution of participants for Participation and Gameplay conditions

Participation	Gameplay		
	VG n=29	MG n=21	NG n=24
VP n=46	18	15	13
MP n=28	11	6	11

The test groups did not differ significantly in terms of sex, age, motivation prior to the experiment and number of gamers. The participants' prior knowledge on aviation and meteorology was tested with a set of five questions in the pre-experiment questionnaire and, after the test, participants were asked to indicate how much prior knowledge they had before the training. The average score on the prior knowledge assessment was 51.9 (SD = 15.0), and the average self-reported score on prior knowledge was 2.5 (SD = 2.1). These scores did not differ significantly between test groups. Furthermore, the two participation groups did not differ in terms of sex, age, motivation prior to the experiment and number of gamers either.

4.4. Adaptations to the Materials

For the second experiment the materials from the first experiment were slightly adapted to accommodate the introduction of the control group (NG), as well as improve their usability.

4.4.1. Website and CloudAtlas game

The website through which the materials were administered, was adapted to allow participants to be assigned to the NG control group. Furthermore, some changes were made to allow the use of Internet Explorer to play the game. No changes were made to the game itself.

4.4.2. Written materials

The written materials were slightly rearranged for an improved presentation. Also, learning objectives were added and the navigation through the pages was made more evident, in order to give participants more control over their learning process. No new content was added to the written materials.

4.4.3. Test and questionnaires

Prior knowledge assessment and test. Five questions about the cloud types and the associated risks were added to the pre-experiment questionnaire, in order to establish an objective measurement of prior knowledge, in addition to the self-report prior knowledge in the post-experiment questionnaire. The actual test remained unchanged, except for a few minor textual corrections.

Pre-experiment questionnaire. A set of questions was added to the pre-experiment questionnaire, relating to the content of the training to ascertain prior knowledge. Also some questions were modified to reflect the changes in the experimental design.

Post-experiment questionnaire. First, a post-experiment questionnaire was added for the NG control group that was included in the second experiment. Second, the post-experiment questionnaire for all participants was expanded with a set of twenty questions (Table 5) based on the Intrinsic Motivation Inventory (IMI) [24, 25].

The IMI is a multidimensional instrument to measure a participant's subjective experience with regard to the activity in an experiment. The current version of the IMI instrument offers subscales to assess a participant's interest/enjoyment, perceived competence, effort, value/usefulness, pressure/tension and perceived choice. Intrinsic motivation is considered to be measured with the interest/enjoyment subscale. Perceived choice and perceived competence are thought to be positive predictors of intrinsic motivation, while pressure/tension is a negative predictor. Relevant items from the IMI subscales can be selected and modified to fit specific activities. Finally, an additional set of four questions relating to the IMI subscales of value/usefulness and pressure was added to the post-experiment questionnaire for the VG group.

Table 5. IMI subscales and items used in post-experiment questionnaires

Interest/Enjoyment
This task was fun to do
I thought this was a boring task (R)
This task did not hold my attention at all (R)
I would describe this task as very interesting
I thought this task was quite enjoyable
Effort/Importance
I put a lot of effort into this task
I didn't try very hard to do well at this task (R)
I tried very hard on this task
It was important to me to do well at this task
I didn't put much energy into this task (R)
Pressure/Tension
I was very relaxed in doing this task (R)
I was anxious while working on this task
I felt pressured while doing this task
* I felt like I was expected to play the game (R)
* I believed I had a free choice about playing the game
Perceived Choice
I felt like it was not my own choice to do this task (R)
I felt like I had to do this (R)
I did this task because I wanted to

Value/Usefulness
I believe this task could be of some value to me
I think that doing this task is useful for training airline pilots
I think this is an important task
* I believed playing the game could be beneficial to me
* I thought playing the game was an important activity
Perceived Competence
I am satisfied with my performance at this task
<i>Note.</i> (R) indicates that the item score is the reverse of the participant's response on that item.
* indicates the item was presented only to participants in the Voluntary Gaming condition.

4.5. Results

A total of 74 participants completed the experiment by using the presented learning materials, taking the test and answering the pre- and post-experiment questionnaires. The game was part of the learning materials of 50 participants (29 VG, 21 MG), of whom 37 played the game. Thirteen VG participants chose not to play the game. Table 6 shows the means and standard deviations on game score, test score and time spent playing the game.

Table 6. Means and SD on results for Participation and Gameplay conditions

Participation		Gameplay						Total n=74	
		VG n=29		MG n=21		NG n=24			
		M	SD	M	SD	M	SD	M	SD
VP n=46		n=18		n=15		n=13		n=46	
	Time played (min)	1.7	1.4	12.9	3.1	N/A	N/A	4.9	6.0
	Game score	659	609	2350	738	N/A	N/A	1024	1119
	Test score	51.0	11.7	50.0	16.9	46.9	7.0	49.5	12.5
	Test - Knowledge	55.8	14.2	54.2	14.8	61.2	10.2	56.8	13.4
	Test - Application	42.5	11.8	48.4	17.4	49.3	15.1	46.3	14.7
MP n=28		n=11		n=6		n=11		n=28	
	Time played (min)	1.2	3.4	14.6	5.8	N/A	N/A	3.6	6.7
	Game score	244	593	2687	952	N/A	N/A	672	1208
	Test score	51.8	9.2	55.1	10.4	44.8	13.1	49.7	11.5
	Test - Knowledge	69.7	15.8	48.3	21.7	52.3	19.8	58.3	204
	Test - Application	46.4	12.1	42.7	9.3	43.3	14.1	44.4	12.1
Total n=74		n=29		n=21		n=24			
	Time played (min)	1.5	2.3	13.4	4.1	N/A	N/A		
	Game score	502	627	2446	795	N/A	N/A		
	Test score	51.3	10.7	51.4	15.2	45.9	10.1		
	Test - Knowledge	61.1	16.1	52.5	16.7	57.1	15.7		
	Test - Application	44.0	11.8	46.8	15.5	46.5	14.6		

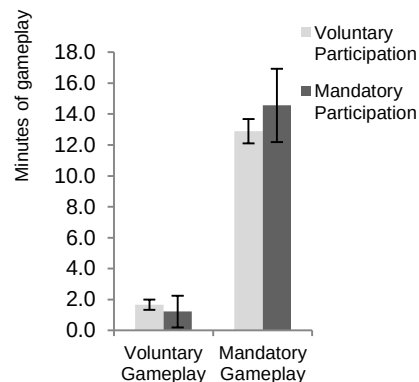


Figure 4. Means and standard errors for time spent playing the game

Gameplay length varied widely as 13 participants in the VG group chose not to play. In the MG group, 5 participants played for more than 15 minutes. Fig. 4 shows that MG players on average spend 11.9 minutes more playing the game than VG players. This difference can be explained by the required minimum of 10 minutes of gameplay for the MG group. On average MG players play significantly longer beyond the required time, than the VG players play in general, $F(1,2) = 9.80$, $p < .001$.

Game scores ranged from 254 to 4060 and overall test scores ranged from 20 to 76. For a more detailed analysis, the overall test results of all participants were split into a score for knowledge questions and one for application questions. Means and standard errors of test scores are visualized in Fig. 5. Two-way analysis of variance shows a significant effect of the interaction of participation and test group on the test score for knowledge questions, $F(2,73) = 3.97$, $p < .01$.

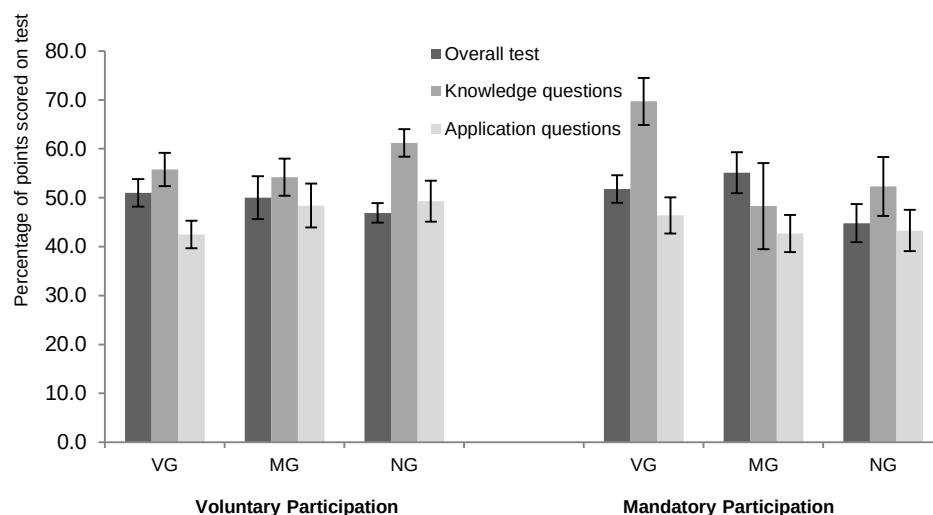


Figure 5. Means and standard errors for test scores in VP and MP conditions

Although the MG participants in the MP group spent more time playing the game than participants in the VP group, this difference is not significant. The participants that played the game, rated their enjoyment on a scale of ten ($N = 35$, $M = 6.3$, $SD = 2.2$). Table 7 shows the means and standard deviations for the enjoyment scores.

Table 7. Means and SD on enjoyment for Participation and Gameplay conditions

Participation		Gameplay					
		VG n=29		MG n=21		NG n=24	
		M	SD	M	SD	M	SD
VP n=46	Game played	n=14		n=15		n=0	
	Enjoyment scored	n=12		n=15		N/A	
	Total	n=18		n=15		n=13	
	Enjoyment	5.4	2.4	6.5	2.0	N/A	N/A
MP n=28	Game played	n=2		n=6		n=0	
	Enjoyment scored	n=2		n=6		N/A	
	Total	n=11		n=6		n=11	
	Enjoyment	7.0	2.8	7.2	2.0	N/A	N/A
Total n=74	Game played	n=16		n=21		n=0	
	Enjoyment scored	n=14		n=21		N/A	
	Total	n=29		n=21		n=24	
	Enjoyment	5.6	2.4	6.7	2.0	N/A	N/A

As in the first experiment, we found that MG participants enjoy the game as much as VG participants (Fig. 6). The same is true for MP participants. They indicate as much enjoyment in the

game as the VP participants. In fact, both the MG and the MP groups give a higher score for enjoyment, although this is not a significant difference.

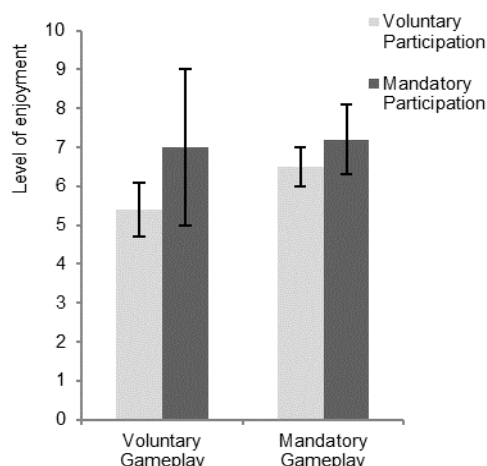


Figure 6. Means and Standard Error for level of enjoyment

MG participants were asked how they felt about having to play the game for at least 10 minutes; bad, neutral or good. Overall, MG participants were a bit negative in the experiment ($N = 21$, $M = 1.7$, $SD = .6$). No difference was found between the voluntary (VP) and mandatory (MP) participants.

In the post-experiment questionnaire, MG participants were asked if they would play the game voluntarily. The NG participants were asked a similar question, after they were informed that other participants had played a game. These questions offered four choices: (1) no, (2) probably not, (3) probably yes, and (4) yes. Table 8 shows the percentages of participants expressing they would play the game for each condition of the experimental design. Table 8 reveals that fewer VG participants who were obligated to participate (MP) chose to play the game than VG participants who volunteered to participate (VP), $t(22.2) = 3.77$, $p < .01$.

An independent samples t-test shows that the willingness of NG participants does not differ significantly from that of the MG participants, $t(38.7) = -1.58$, $p = .12$. Nor is there a significant difference between the percentage of NG participants willing to play and the percentage of VG participants that actually played, $t(51.0) = -1.90$, $p = .06$.

Table 8. Reported willingness to play the game voluntarily versus actual gameplay

Participation	Gameplay					
	Reported willingness				Actual gameplay	
	MG n=21		NG n=24		VG n=29	
VP n=48	n=15	67%	n=13	77%	n=18	78%
MP n=26	n=6	33%	n=9	82%	n=11	18%
Total	n=21	57%	n=24	79%	n=29	55%

Motivation is one of the main factors in this research. In the pre-experiment questionnaire, participants were asked to indicate their motivation to participate in the experiment on a scale of ten. Table 9 shows the mean and standard deviations.

F-tests revealed a significant difference. The motivation reported by VP participants is significantly higher than that of MP participants, $F(1,72) = 12.52$, $p < .001$. This is in accordance with the general expectation that volunteers are more motivated.

Table 9. Self-reported levels of motivation separated for the Participation conditions

Measure	Participation			
	VP n=46		MP n=28	
	M	SD	M	SD
Motivation	7.4	1.1	6.0	2.4

The post-experiment questionnaires contained a set of twenty questions in regard to motivation, based on the IMI scales [24, 25]. Two-way ANOVA tests were conducted to examine the influence of Participation and Test Group on the six IMI subscales. The significance levels are shown in Table 10.

A simple main effect showed that VP participants reported to experience significantly more choice than MP participants ($p < .01$). No interaction effects or other main simple effects have been found.

Table 10. Significance levels of Between Subject Effects for the IMI subscales

IMI Subscale	Between Subject Effects		
	Gameplay (VG, MG, NG)	Participation (VP, MP)	Interaction of Gameplay and Participation
Interest	.98	.28	.68
Pressure	.43	.16	.08
Choice	.28	.00*	.14
Value	.58	.88	.06
Effort	.91	.78	.28
Competence	.69	.17	.43

Note. * $p < .01$

4.6. Discussion

There are six aspects of interest which lead to a discussion. These are test scores, gameplay, motivation, incentive, reward and cultural aspects.

Test scores. The study sought to investigate the impact of voluntariness on the learning effect of a serious game. The learning effect of the serious game was measured by a test, taken shortly after the training. The data showed that mandatory players (MG) spent more time playing the game. However, performance on the test does not differ statistically between the test groups. Therefore, time spend on training does not appear to be a factor. The topic of the training and the game may have been outside the area of interest for most participants, or the game may not be as effective as expected.

Gameplay. The second aspect of interest was gameplay, measured in game score and duration. Almost half of the voluntary players (VG) chose not to play the game at all and the participants who did play, played for a shorter period of time than mandatory players (MG). This difference can be explained by the required minimum of 10 minutes of gameplay for the MG group. Only one VG participant played for more than 10 minutes, all others played less than 4 minutes. Even after correcting for the required minimum of 10 minutes, MG participants played significantly longer. This may indicate that MG participants become more engaged in the game than VG participants. This seems to be unrelated to the fun of the game, as participants in the VG and MG groups report an equal score on enjoyment.

MG players from the mandatory participation (MP) group played, almost two minutes longer than MG players from the voluntary participation (VP) group; an interesting although not significant difference. No explanation has been found for this difference, but it could be the effect of the obligation toward the lecturer. On the other hand, the VG players in the MP group play half a minute shorter than the VG players in the VP group, on a total gameplay time of two minutes. It seems that, contrary to the MG players, VG players do not feel obligated to keep playing.

Motivation. We expected freedom of choice to motivate voluntary players and encourage them to accomplish better results. However, voluntary players (VG) did not do better in the game and on the test, nor did they score higher on the level of enjoyment than mandatory players (MG), irrespective of whether they participated voluntarily (VP) or as part of a homework assignment (MP).

Additionally, it would be understandable for a player to experience negative feelings when obligated to play the game, but mandatory players (MG) reported a neutral feeling about having to play the game for a minimum amount of time. The data suggest that a negative feeling about the minimum time is unrelated to voluntary or mandatory participation, nor is it related to a participant's motivation to participate in the experiment.

The majority of both mandatory players (MG) and control group participants (NG), indicated that they would play the game if they were offered a choice. The percentage of MG and NG participants, who said they would play the game without the obligation, was higher than the percentage of voluntary players (VG) who actually did play the game.

Incentive. We found that VP participants were more motivated than MP participants. A possible explanation for this is that participants with low motivation may not participate if they have the choice (as is the case in VP), but in MP, they are obligated to participate, which in effect may reduce the overall score for motivation.

Reward. In the experiment a number of different rewards were offered to students for their participation. Similar to the first experiment, voluntary participants (VP) who were invited to participate by their teacher, were offered a chance of winning a € 100 gift card, while participants through the ERAS network received school credit and were not eligible to win the gift card. Based on the individual lecturer's options, some mandatory participants (MP) were offered the chance of winning the gift card, some received school credit and others did not get any reward. The distribution of rewards is shown in Table 11. These rewards may have influenced students' motivation.

Table 11. Distribution of reward for Participation and Gameplay conditions

Participation		Gameplay		
		VG n=29	MG n=21	NG n=24
VP n=46	School credit	n=0	n=0	n=0
	Chance of winning	n=18	n=15	n=13
	No reward	n=0	n=0	n=0
	Total	n=18	n=15	n=13
MP n=28	School credit	n=2	n=0	n=2
	Chance of winning	n=3	n=3	n=5
	No reward	n=6	n=3	n=4
	Total	n=11	n=6	n=11

Cultural aspects. The participants in the voluntary participation group (VP) were all enrolled in schools in the Netherlands, whereas the students in the mandatory participation group (MP) come from schools in the Netherlands, Thailand and Lebanon. As these countries have different cultures, cultural differences may have had an impact on the outcomes of the experiment. Among other things, students in Thailand and Lebanon may have a more formal relationship with their teachers than Dutch students, possibly resulting in a stronger sense of obligation. Table 12 shows the distribution of the different origins of participants.

Table 12. Distribution of origins for Participation and Gameplay conditions

Participation		Gameplay		
		VG n=29	MG n=21	NG n=24
VP n=46	the Netherlands	n=18	n=21	n=13
	Thailand	n=0	n=0	n=0
	Lebanon	n=0	n=0	n=0
	Total	n=18	n=21	n=13
MP n=28	the Netherlands	n=6	n=3	n=4
	Thailand	n=3	n=1	n=5
	Lebanon	n=2	n=2	n=2
	Total	n=11	n=6	n=11

Note. 11% of the students enrolled in 2014 to (applied) universities in the Netherlands are international students.

5. General discussion

Based on the two experiments, it appears that we may conclude that participants with a stronger obligation played as well or better than participants who were free to play the game. Also, they enjoyed the game as much or more. Surprisingly, it may be concluded that a little force increases the time spent in the game, which in turn may improve the learning effect.

The two experiments described in the article are similar, where the second experiment is designed to dig deeper into the unexpected results from the first experiment. Below we discuss settings and findings. The main difference was between the informal setting of the first experiment and the formal educational setting of the second experiment. Additionally, the second experiment had a larger sample size, combining voluntary participants with mandatory participants, who have an obligation toward their teacher or school.

The CloudAtlas game that was used in both experiments, was developed for this research. The findings indicate that players may need some time to get engaged in the game. This may have to do with the quality, playability, difficulty or even graphical design of the game. In contrast to the first experiment, gamers and non-gamers did not differ significantly in the second experiment.

In both experiments (part of) the participants volunteered to participate. This in itself may affect the motivation of the participants. For most of the participants, the topic of the training and the game was irrelevant for their studies or everyday life. This may lead to a lower motivation [26].

As the learning content relates to knowledge and interpretation of clouds for aviation, the actual target groups of the serious game are pilot or meteorology trainees. None of the participants belonged to the actual training group, but the level of education is similar. The actual target group may be more motivated to learn about the topic, but otherwise the differences are negligible.

Motivation plays an important role in this experiment, both intrinsic and extrinsic. A meta-analysis of studies on the effect of extrinsic rewards on intrinsic motivation has shown that tangible rewards for interesting tasks may have a negative effect on intrinsic motivation [27].

Participants of the MG group continue to play longer than VG participants. The exact cause of this effect was not established in the experiments, and may require further research. Possibly there is a threshold before getting engaged in the game, that is crossed by MG players as they have a required gameplay.

In both experiments, having a minimum amount of required gameplay was proven not to have a negative effect on the enjoyment of the game. The second experiment confirmed that setting a minimum leads to longer gameplay and higher test scores. Optimal instruction may require testing different durations of mandatory gameplay, to find the minimum period of time that motivates students to play longer.

As our results contradict much of the existing, but almost exclusively non-empirical, literature on voluntary play, an obvious avenue for future research is to explore why our findings differ. Are there certain conditions in which voluntary is more effective than mandatory gameplay? To what extent are the outcomes affected by factors such as motivation, cultural aspects or level of education?

As voluntary play is thought to be indispensable for gaming and thus for serious gaming, using a serious game in a training curriculum would disadvantage students who do not wish to play the game. An alternative method would have to be offered in order to give them a fair chance of successfully completing the course, leading to higher training expenses. Even when taking the shortcomings of the study into account, these two experiments indicate that mandatory play of a serious game does not negatively impact the learning outcome or the player experience. If this effect holds up in continued research, it may be concluded that adding a serious game to a professional training curriculum does not require an alternative training method.

6. Conclusion

Both experiments aimed to determine whether, and to what extent, gameplay and learning effect of a serious game are affected by the freedom to choose to play or not to play the game. In an ideal situation, this effect would have been found and nothing but the voluntariness would have had an impact on it. However, we found no learning effect of the game at all. We did find an interesting effect regarding gameplay and enjoyment.

We expected that using the game voluntarily as a learning tool, would result in improved player performance in a test, in comparison to the results after mandatory gameplay. This result was found in neither the first nor the second experiment.

Our findings indicate that participants with a stronger obligation played at least as long and as well as participants who were free to play the game. Also, they enjoyed the game as much or even more. Furthermore, we found that a little force increases the time spent in the game, which in turn may improve the learning effect.

The surprising finding, that mandatory gameplay in the game does not appear to ruin the enjoyment and engagement in the game, challenges the assumption of many game design theorists and practitioners that games need to be played voluntarily in order to be engaging, fun, and effective. The studies discussed in the article indicate that mandatory gameplay is just as much fun as voluntary gameplay. Follow-up studies with a homogenous group of participants and no disturbing reward mechanisms are required to determine if this indication holds up, and what the reasons are.

References

- [1] Michael, D., Chen, S., *Serious Games: Games That Educate, Train and Inform*. Thomson, Boston (2006).
- [2] Wu, W., Hsiao, H., Wu, P., Lin, C., Huang, S., Investigating the Learning-Theory Foundations of Game-Based Learning: a Meta-Analysis. *Journal of Computer Assisted Learning* 28(3), 265-279 (2011). <http://dx.doi.org/10.1111/j.1365-2729.2011.00437.x>
- [3] Wouters, P., Van Nimwegen, C., Van Oostendorp, H., Van Der Spek, E., Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. *Journal of Educational Psychology* 105(2), 249-265 (2013). <http://dx.doi.org/10.1037/a0031311>
- [4] Heeter, C., Lee, Y., Magerko, B., Medler, B., Impacts of Forced Serious Game Play on Vulnerable Subgroups. *International Journal of Gaming and Computer-Mediated Simulations* 3(3), 34-53 (2011). <http://dx.doi.org/10.4018/jgcms.2011070103>
- [5] Mollick, E., Rothbard, N., *Mandatory Fun: Consent, Gamification and the Impact of Games at Work*. The Wharton School Research Paper Series (2014).
- [6] Susi, T., Johanneson, M., Backlund, P., *Serious Games - An overview*. Technical Paper, University of Skövde, Skövde (2007).
- [7] Djaouti, D., Alvarez, J., Jessel, J., Rampnoux, O., Origins of Serious Games. In Ma, M., Oikonomou, A., Jain, L. (eds.), *Serious Games and Edutainment Applications*. Springer-Verlag, London (2011) 25-44. http://dx.doi.org/10.1007/978-1-4471-2161-9_3
- [8] Abt, C., *Serious Games*. Viking Press, New York (1970).
- [9] Sawyer, B., Rajeski, D., *Serious Games: Improving Public Policy through Game-Based Learning and Simulation*. Woodrow Wilson International Center for Scholars, Washington D.C. (2002).
- [10] Salen, K., Zimmerman, E., *Rules of Play: Game Design Fundamentals*. MIT Press, Cambridge (2004).
- [11] McGonigal, J., *Reality Is Broken*. Penguin Press, New York (2011).
- [12] Huizinga, J., *Homo Ludens: A Study of the Play Element in Culture*. Beacon Press, Boston (1955).
- [13] Caillois, R., *Man, Play and Games*. University of Illinois Press, Champaign (1961).
- [14] Garris, R., Ahlers, R., Driskell, J., Games, Motivation, and Learning: a Research and Practice Model. *Simulation & Gaming* 33(4), 441-467 (2002). <http://dx.doi.org/10.1177/1046878102238607>
- [15] Prensky, M., *Computer Games and Learning: Digital Game-Based Learning*. In Raessens, J., Goldstein, J. (eds.): *Handbook of Computer Game Studies*. MIT Press, Cambridge (2005) 97-122.
- [16] Breuer, J., Bente, G., Why So Serious? On the Relation of Serious Games and Learning. *Eludamos, Journal for Computer Game Culture* 4(1), 7-24 (2010)
- [17] Cain, J., Piascik, P., Are Serious Games a Good Strategy for Pharmacy Education? *American Journal of Pharmaceutical Education*, 79(4), Article 47 (2015). <http://dx.doi.org/10.5688/ajpe79447>
- [18] Fulton, S., Schweitzer, D., Impact of Giving Students a Choice of Homework Assignments in an Introductory Computer Science Class. *International Journal of the Scholarship of Teaching and Learning* 5(1), Article 20 (2011). <http://dx.doi.org/10.20429/ijstl.2011.050120>
- [19] Becker, K., How Much Choice is Too Much? *Inroads* 38(4), 78-82 (2006). <http://dx.doi.org/10.1145/1189136.1189176>

- [20] Gauthier, A., Corrin, M., Jenkinson, J., Exploring the Influence of Game Design on Learning and Voluntary Use in an Online Vascular Anatomy Study Aid. *Computers & Education* 87, 24-34 (2015). <http://dx.doi.org/10.1016/j.compedu.2015.03.017>
- [21] Kuindersma, E. C., Van der Pal, J., Van den Herik, H. J., Plaat, A., Voluntary Play in Serious Games. Paper presented at the GaLA Conference, Rome (2015).
- [22] Hernon, P., Whitman, J. R., Delivering Satisfaction and Service Quality: a Customer-Based Approach for Libraries. American Library Association, Chicago, IL (2001).
- [23] Coelho, P., Esteves, S., The Choice Between a Five-Point and a Ten-Point Scale in the Framework of Customer Satisfaction Measurement. *International Journal of Market Research* 49(3), 313-339 (2007).
- [24] Ryan, R. M., Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology* 43, 450-461 (1982). <http://dx.doi.org/10.1037/0022-3514.43.3.450>
- [25] Selfdeterminationtheory.org: Intrinsic Motivation Inventory (IMI) Scale Description. Retrieved on March 22, 2016 from <http://selfdeterminationtheory.org/intrinsic-motivation-inventory/>.
- [26] Vroom, V. H., Work and motivation. Wiley, New York (1964).
- [27] Deci, E. L., Koestner, R., Ryan, R. M., A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125(6), 627 (1999). <http://dx.doi.org/10.1037/0033-2909.125.6.627>