Does Gamification of a Student Response System Boost Student Engagement, Motivation and Learning?

- An Evaluation of the Game-based Student Response System Kahoot!

Alf Inge Wang[1], Meng Zhu and Rune Sætre,
Dept. of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT
The article describes an experiment where the game-based student response system Kahoot! was compared to a traditional non-gamified student response system, as well as the usage of paper forms for formative assessment. The goal of the experiment was to investigate whether gamified formative assessments improve the students' engagement, motivation, and learning. In the experiment, the three different formative assessment tools/methods were used to review and summarize the same topic in three parallel lectures in an IT introductory course. Pre- and a post-test were used to assess the learning outcome of the lectures, and a questionnaire was used to get data on the students' engagement and motivation. The results show significant improvement in motivation and engagement for the gamified approach, but we did not find significant learning improvement. Further, the students that used the traditional student response system were only slightly more (not significant) motivated and engaged compared to the students that used of paper forms. We also found higher variation in motivation and engagement related to gender and whether students played games or not for formative assessment using paper forms compared to the two other tools. Female students and non-gaming students were significantly more negative towards using paper forms for formative assessment compared to male and gaming students respectively. The gamified approach was least sensitive to gender and whether students play games.

Keywords: Game-based learning; Student response system; Student engagement; evaluation; formative assessment

1. Introduction

Prototypes of student response systems (SRSs) have been around since the sixties (Judson 2002), and these systems started to be used in biology and chemistry teaching in the early seventies (Bessler and Nisbet 1971, Casanova 1971). The first generation of SRSs was based on special hardware that allowed the students to give their answers using clickers, key-pads, handsets or zappers (Caldwell 2007). A major disadvantage with this first generation of systems was that they required investment in hardware devices and infrastructure as well as administration and maintenance of the hardware and software. The Bring Your Own Device wave has opened up for a new generation of SRSs, where students can use their own devices to respond. After the introduction of smart phones and tablets providing support for easy wireless network access and support for HTML5, many new SRSs and similar tools have populated the market: for example Socrative (Coca and Slisko 2013), Quizlet (Gruenstein, McGraw et al. 2009), Poll Everywhere (Sellar 2011), iClicker (Lucas 2009), and Learning Catalytics (Schell, Lukoff et al. 2013). The use of HTML5 web-technology makes it possible to use these systems without installing any applications, and opens up for a range of new ways of interaction in the classroom. Kahoot! is a game-based SRS (GSRS) that was introduced to the public in the fall 2013. The main difference between a GSRS and a SRS is that the game-based version focuses more on engaging and motivating the students through

[1] Corresponding author: alfw@idi.ntnu.no , +47 7359 4485, IDI/NTNU, Sem Sælands vei. 7-9, NO-7491 Trondheim, Norway

attractive graphical user-interfaces and audio, as well by gamifying the whole student response experience. The gamification is done by temporarily transforming the classroom into a game show as shown on TV, where the teacher play the role of a game show host and the students are the competitors. Well-designed video games are said to be learning machines (Gee 2003), and they have the potential to get the players so motivated and engaged that they are not aware that learning is actually happening. In K-12[2], games have been found to be beneficial for academic achievement, motivation and classroom dynamics (Rosas, Nussbaum et al. 2003). Games have also been found to have a similar effect in higher education (Sharples 2000). Previous research indicates that games can be made an integrated part of traditional classroom lectures to improve learning, motivation and engagement (Carver Jr, Howard et al. 1999, Carnevale 2005, Wang, Øfsdal et al. 2007, Wang, Øfsdal et al. 2008, Wu, Wang et al. 2011).

This article presents an experiment where the game-based SRS Kahoot! was compared to a simple traditional SRS, as well as to the usage of paper forms for formative assessment. The goal of the experiment was to investigate whether adding gamification to formative assessment could improve the students' engagement, motivation and learning. The rest of the article is organized as follows. Section 2 presents material and methods including related work, a description of the formative assessment tools used in the experiment, the data sources used, the research context and participants of the experiment, the experiment procedures, and the data analysis. Section 3 presents the results from the experiment. Section 4 discusses the results and their validity of the results. Section 5 concludes the article.

2. Material and Methods

The goal of the controlled experiment presented in this article was to investigate if the choice of method for doing a formative assessment affects the student engagement, motivation and learning. This section presents the related work, the three assessment tools used, the data sources, the research context and participants, research procedures, and the method for data analysis.

2.1 Related Work

This section describes other experiments using SRS or games for learning in classrooms. Game-based SRS is not covered as we did not find any existing literature on this topic.

There have been many experiments and studies conducted on SRSs, and a literature study from 2007 reports that such systems have been found to have a positive effect on student exam-performance, and that they create a more positive and active atmosphere in classrooms (Caldwell 2007). More specifically, students using SRSs were twice as likely to work on a problem presented during class (Cutts, Kennedy et al. 2004), student attendance rose to 80-90% (Burnstein and Lederman 2001), and about 88% of the students either "frequently" or "always" enjoyed using the SRS in class (Caldwell 2007). Further, Caldwell's survey summarized some common uses of clicker questions found in the literature: to increase or manage interaction, to assess student preparation and ensure accountability, to find out more about students, for formative assessment, for quizzes or tests, to do practice problems, to guide thinking review or teach, to conduct experiments, to make lectures fun, to differentiate instruction, and to prompt discussion. Another study summarizes similar findings of benefits of SRSs in the three areas Classroom environment, Learning, and Assessment (Kay and LeSage 2009). SRSs were found to improve attendance, provide more focused students, provide anonymous student participation, improved student engagement, increase learning performance, improved teaching, and generally improve interaction between teacher and students. The benefits listed in these surveys are all benefits we have experienced from using game-based SRS in classrooms as well.

[2] Denotes the sum of primary and secondary education in the US and other countries

There are several studies on the effects of educational games related to learning outcome and increased motivation, and we will present some of these studies here. One study explored the impact of using a game named Supercharged! on pre-service teachers' understanding of electromagnetic concepts compared to students who conducted a more traditional inquiry oriented investigation of the same concept (Anderson and Barnett 2011). The effectiveness was investigated through an experiment that used both qualitative and quantitative data that included pre- and post-scores, student notebooks, video recordings of laboratory activities and observations. The results of this study showed that the group using the video games outperformed the group that did not use the video game in terms of learning outcomes (statistically significant). In another study, the video game *relative asteroids* was used to teach physics (Carr and Bossomaier 2011). A pre- and post-test with 8 questions was used to measure the learning outcome. The results showed improvement of the test scores for new learners. The students also found that learning physics through a game was motivating and engaging, and it was effective at improving their comprehension of physics. In a chemistry class, an experiment was conducted to compare students' achievement and attitude from traditional vs. game-based teaching methods (Tüysüz 2009). The statistically significant conclusion of the study was that game-based learning increased the students' achievement in chemistry compared to traditional learning methods. The study also showed that game-based learning increased the students' interest in the course: they enjoyed the course more, and were more focused and engaged on the subject being taught. Similar results on improved learning outcome were also found for using a web-based adventure game to teach neuroscience (Miller, Schweingruber et al. 2002), in an experiment comparing teaching computer memory knowledge with a game vs. a non-game application (Papastergiou 2009), and for using a mobile game to engage students in arithmetic practices (Liao, Chen et al. 2011). In the computer memory study, the students that used a game found this learning approach significantly more appealing and educational fruitful than the students with the non-game application (Papastergiou 2009). The same students also found their learning approach more engaging, more effective, more active and relaxed compared to the students that used the non-game approach. There are also studies that show that introducing games into the classroom not always produce positive results and can result in complaining students and lack of motivation (Squire 2005).

Kahoot! represents a new generation of student-response systems that has a main focus on student motivation and engagement through gamification. The tool is a result of the research project Lecture Quiz that started in 2006 (Wang, Øfsdal et al. 2007), where results from experimentation of early prototypes showed positive results in terms of increased engagement, motivation and perceived learning (Wang, Øfsdal et al. 2008, Wu, Wang et al. 2011). Educational games compared to mainstream entertainment games are known to suffer from running on very few platforms (usually Windows PCs), too simplistic, being single player and offline, offering low production value, and are typically more targeted towards parents, teachers and formal learning curriculum than being fun for the students (Kirriemuir and McFarlane 2004). This is especially true when educational games try to copy existing game concepts and add some learning on top of it. Kahoot! was not designed to copy any existing game, but rather to find a game concept that could fit a classroom setting and that could be alignment with Tom Malone's theory of intrinsically motivating instructions (Malone 1980). Malone's theory lists three categories that make things fun to learn: Challenge (goals with uncertain outcomes), Fantasy (captivate through intrinsic or extrinsic fantasy), and Curiosity (sensor curiosity through graphics and audio, and cognitive curiosity). As the game should be used in the classroom, it was also important to incorporate a social game play. The result was to develop a game concept where the classroom temporarily is changed to a game show where the teacher is the game host and the students are the competitors. The *challenge* is to answer questions and compete against other players, the *fantasy* is being a part of a game show, and the *curiosity* is provided through inspiring graphics and audio, as well as solving a cognitive puzzle. The lack of variety in game play is compensated by the competitive nature of playing against a whole class of students. Reports

from happy teachers and students all over the world give an indication that the concept works as intended. Learning games are commonly used to review facts using multiple-choice questions similar to what is done in Kahoot!. However, games can also be used to teach skills, judgment, behaviors, theories, reasoning, process, procedures, creativity, language, systems, observation, and communication using various approaches (Prensky 2005).

2.2 Three Methods for Running Quizzes in a Classroom

In the experiment presented in this article, three different approaches were used to run a quiz as a part of a lecture. The quiz was used to review what being taught in the lecture and consisted of 12 multiple-choice questions. The quiz methods used were paper forms, a simple student response system named Clicker, and a game-based student response system named Kahoot!. Figure 1 shows pictures from the three lectures of students doing the quiz (left: paper quiz, middle: Clicker, right: Kahoot!). The three quiz methods will now be described more in detail.

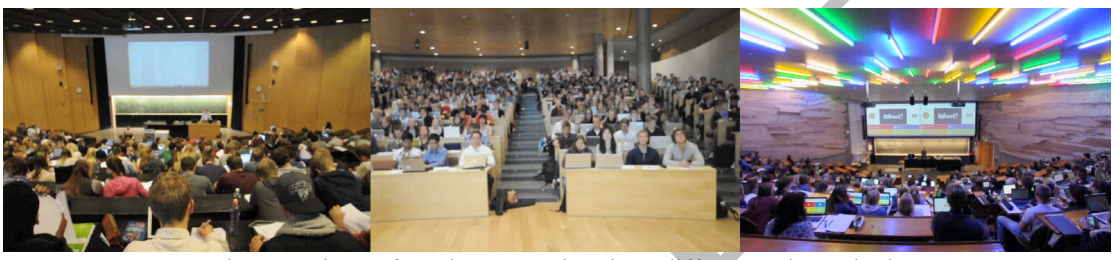


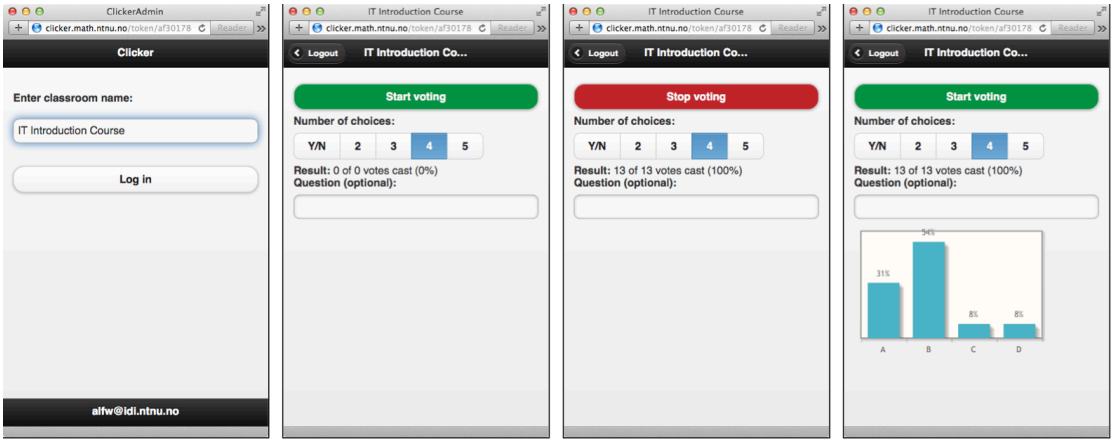
Figure 1 Picture from lectures using three different quiz methods

2.2.1 Paper Form and Hand Raising

The paper quiz is an analogue, well-known and proven approach for running quizzes in a classroom. Before a lecture, the teacher prepares paper forms with multiple-choice questions where students can tick of one of two to four answers using a pen or a pencil. During the lecture, the teacher hands out the paper forms, and the students answers as well as they can. The normal procedure for such quizzes is that the teacher then collects the paper forms, and the students get feedback in the following lecture on how they have performed. In our experiment, we changed the review part to be more compatible with the digital counter-parts. After the students had completed their forms, the teacher reviewed their answers by going through the questions asking students to raise their arms for the alternative they had answered. In this way, the teacher got to know how the students had answered and gave him the chance to give feedback to the class. The paper forms used consisted of a quiz of twelve questions, all with four alternative answers.

2.2.2 The Clicker Student Response System

Clicker is a simple student response system (SRS) allowing students to give their responses to questions being asked using a web-browser on any digital device. The questions and answers are typically shown using another tool like Keynote, Prezi or PowerPoint, and Clicker is used to collect "votes" from the students. Figure 2 shows an overview of the Clicker SRS and the steps needed to get responses from students. First, the teacher needs to name a classroom which will be the ID used for students to connect (Figure 2a). Second, the teacher chooses how the students can response to a question (Figure 2b). Third, the teacher has to show the questions and answers on the screen using PowerPoint or similar tool as well as he starts the voting process (Figure 2c). In this step, the teacher can monitor how many students that have given their answers. Forth, the teacher stops the voting, and a distribution of how the students have voted is shown (Figure 2d). The distribution of answers is also shown on the student client. The Clicker SRS does not directly give the teacher or students any feedback on correctness of answers. It is up to the teacher to comment on correctness of the students' responses based on the given distribution of votes.



a) Enter classroom b) Choose choices c) Collect votes d) View results

Figure 2 Overview of the Clicker Student Response System

2.2.3 The Kahoot! Game-based Student Response System

Kahoot! is a game-based student response system (GSRS) launched by the teacher in a web-browser on a laptop connected to a large screen. Unlike the Clicker SRS, Kahoot! provides a tool for creating quizzes including adding pictures and YouTube videos to the questions. Kahoot! also makes it possible to publish and share your own quizzes, and edit quizzes made by others. Another difference is the way Kahoot! is played. Students will log into the system using a game pin (a number) and a nickname. The goal for the students is to answer the correct answer as fast as possible to get as many points as possible. Figure 3 shows how Kahoot! is played. A question is shown on the large screen along with four or less alternatives in different colors with associated graphical symbols. The students give their answers by choosing the color and symbol she or he believes corresponds to the correct answer.

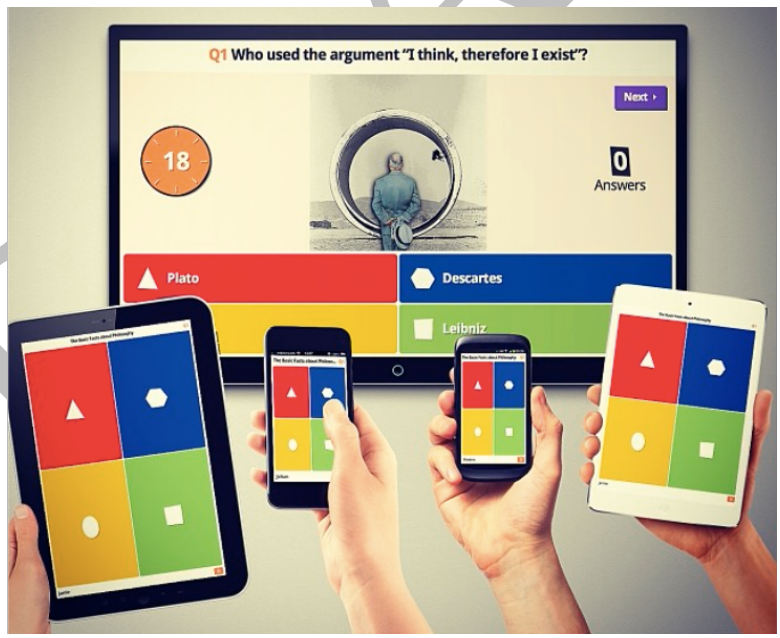


Figure 3 Playing Kahoot!

Between every question, a distribution of how the students answered is shown before a scoreboard of the five best players. The students get individual feedback on their questions in terms of correctness, the number of points, the ranking, how far the student is behind the student ranked above, and the correct answer if wrong answer is given. At the end of a Kahoot! session, the winner's nickname and points will be shown on the large screen. During the quiz, Kahoot! uses a playful graphical user interface as well as music and sounds to give it a playful and competitive atmosphere similar to a game show on TV. The students are also

asked to give feedback on the quiz they have played through giving scores on whether the quiz was fun, educational, can be recommended to others, and how you generally feel about the quiz. Finally, Kahoot! provides the functionality for the teacher to download the results from the quiz in an Excel spreadsheet.

2.3 Data Sources

The instruments for collecting data in our experiment included a domain knowledge test and a questionnaire on students' engagement and motivation. The domain knowledge test consisted of seven multiple-choice questions developed by a domain expert, and this questionnaire was used both as a pre-test and a post-test to measure the knowledge before and after the lecture.

The motivation questionnaire was developed to measure the motivation and the engagement of the students. The questionnaire was adapted from the course motivation survey (CMS) (Kebritchi, Hirumi et al. 2010) to our research context, and integrated with relevant questions in the Motivated Strategies for Learning Questionnaire (MSQL) (Pintrich 1991) and (Lepper, Corpus et al. 2005). The questionnaire used a five-point Likert scale from strongly disagree to strongly agree.

2.4 Research Context and Participants

The experiment was performed in the IT introductory course (TDT4105) at Norwegian University of Science and Technology (NTNU). There were two reasons for choosing this particular course for doing the experiment. First, the IT introductory course is a large course with many students, meaning that it would be possible to collect data from many subjects. Second, due to the size of this course, the same lecture has to be taught in three parallels. This means that the same teacher will teach exactly the same lecture for three parallels of students. The IT introductory course is a mandatory course for all first year students at the university, giving that the groups of students in the experiment should be fairly uniform.

The experiment was conducted over three days at the end of September 2013, and the topic of the lecture was on basic computer knowledge.

The distribution of the subjects that completed the questionnaires was:

- Paper quiz: 127 subjects, 58% female and 42% male students, and 50% of the students play video games at least one hour a week.
- Clicker quiz: 175 subjects, 37% female and 63% male student, and 62% of the students play video games at least one hour a week.
- Kahoot! quiz: 82 subjects, 54% female and 46% male students, and 50% of the students play video games at least one hour a week.

2.5 Procedures

The lecture in the experiment was conducted according to Figure 4. First, the teacher introduced the lecture by presenting the agenda and the current topic, before the students carried out a paper pre-test on the lecture's topic. Second, the teacher taught the topic *basic computing* using Power-point slides. Third, a quiz on the topic was carried out in three variations for the three parallels (Paper, Clicker, and Kahoot!). Forth, at the end of the lecture, the students had to fill in a motivation questionnaire as well as doing the paper post-test (same as the pre-test). Fifth, some students were contacted after the lecture for an interview. The only part that was different for the three parallels was the method for running the quiz.

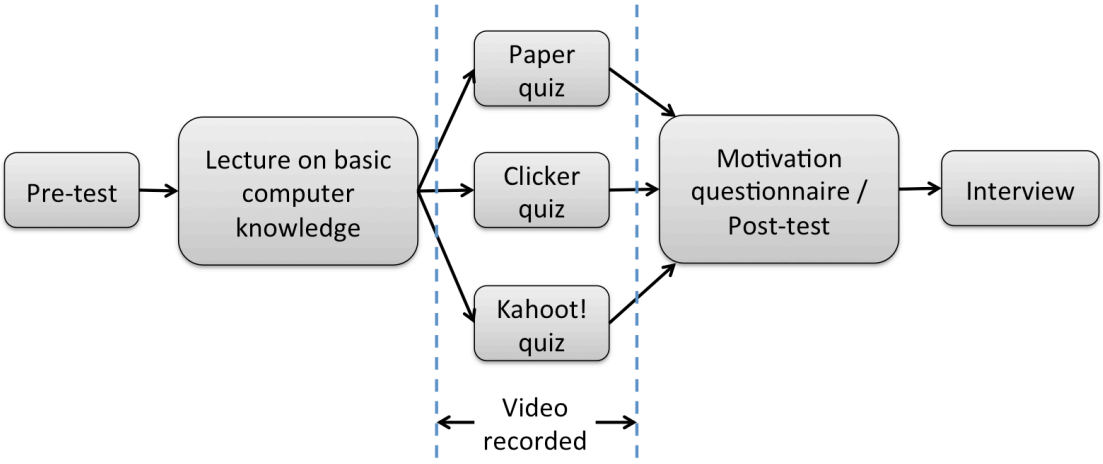


Figure 4 Experiment Procedures

2.6 Data Analysis

The answers from the pre- and post-tests were evaluated to a score from 0 to 7 points, where each correct answer contributed one point to the total score. The learning outcome was computed by comparing the difference between the post- and the pre-test scores. The Mann-Whitney test was used to compare the difference in the learning outcomes from the different quiz methods (Paper, Clicker and Kahoot!). The Mann-Whitney test is a nonparametric test for the significance of the difference between the distributions of two independent samples of difference sizes.

The Kruskal-Wallis test was ran on the data from the motivation questionnaire to investigate the differences between the responses from the three groups Paper, Clicker, and Kahoot!. The Kruskal-Wallis test is a nonparametric test for the significance of the differences among the distributions of in our case three independent samples that had difference sizes. Further we used the Mann-Whitney test for investigating differences in engagement and motivation of the two groups male vs. female students, and the two groups students playing games vs. students not playing games.

3. Results

This section presents the results from the controlled experiment. In the analysis we looked at differences in students engagement and motivation in regards to the use of quiz method, gender and whether the students played games or not. This section also reports on differences in the learning outcome comparing quizzing using paper vs. using Kahoot!. Unfortunately do we not have the data on learning outcome for the Clicker system, as the students did not get the time to complete the post-test for this parallel.

3.1 Comparison of the Use of Paper, Clicker and Kahoot!

Table I shows the descriptive statistics and the results from the Kruskal-Wallis test comparing how the three method of running quizzes (Paper, Clicker and Kahoot!) affects the students engagement and motivation. To improve readability, a 3-point Likert scale is used to show the result (Disagree, Neutral and Agree). This means that the responses Strongly disagree and Disagree are grouped into to the category *Disagree*, and Strongly agree and Agree are grouped into the category *Agree*. The table shows the results from the Kruskal-Wallist test for $k=3$ with $n_{paper}=127$, $n_{clicker}=175$, $n_{kahoot!}=82$, resulting in $df=2$. Statistically significant differences are highlighted in the table.

Table I. Comparison of use of Paper, Clicker and Kahoot!

Statement	Group	Disagree	Neutral	Agree	H	P
1. I concentrated on the quiz to get correct answer	Paper	8%	31%	61%	4.59	0.1008
	Clicker	7%	22%	71%		
	Kahoot!	6%	16%	78%		
2. I wanted to answer quiz without help from others	Paper	20%	20%	60%	5.19	0.0746
	Clicker	10%	17%	73%		
	Kahoot!	9%	20%	72%		
3. I did the quiz only because the teacher told me to	Paper	50%	25%	25%	14.2	**0.0008**
	Clicker	56%	29%	15%		
	Kahoot!	77%	18%	5%		
4. To complete the quiz gave me satisfaction	Paper	35%	33%	32%	3.1	0.2122
	Clicker	28%	41%	31%		
	Kahoot!	22%	38%	40%		
5. I enjoyed the quiz so much that I want to know more about the topic	Paper	37%	43%	20%	1.4	0.4966
	Clicker	34%	39%	27%		
	Kahoot!	35%	49%	16%		
6. The quiz was boring and not engaging	Paper	60%	34%	6%	9.59	**0.0083**
	Clicker	63%	26%	10%		
	Kahoot!	84%	12%	4%		
7. It gave me satisfaction to complete the quiz in a satisfactory way	Paper	18%	41%	41%	1.65	0.4382
	Clicker	16%	40%	44%		
	Kahoot!	15%	34%	51%		
8. I do not think this lecture was worth my time and effort	Paper	56%	29%	15%	2.79	0.2478
	Clicker	51%	32%	17%		
	Kahoot!	63%	24%	12%		
9. To score high on the quiz was the most satisfactory in the lecture	Paper	38%	40%	22%	15.91	**0.0004**
	Clicker	35%	37%	27%		
	Kahoot!	22%	26%	52%		
10. I wished to have a better score on the quiz than most other students in the class	Paper	24%	39%	38%	20.66	**<0.0001**
	Clicker	16%	41%	43%		
	Kahoot!	4%	28%	68%		
11. I felt increased pulse when I answered questions in the quiz	Paper	65%	22%	13%	31.78	**<0.0001**
	Clicker	66%	26%	8%		
	Kahoot!	34%	20%	46%		

Table I shows statistically significant differences in statements 3, 6, 9, 10 and 11 where $p \leq 0.05$. For statement 3, students doing the Kahoot! quiz to a larger extent disagreed (77% vs. 50% for paper and 56% for Clicker) that they did the quiz *only because* the teacher told them. Similarly for statement 6, Kahoot! students to a larger extent disagreed (84% vs. 60% and 63%) that the quiz was boring and not engaging. The analysis also reveals that students doing the Kahoot! quiz thought to a larger degree that the most satisfactory in the lecture was to score high on the quiz (52% vs. 22% and 27%), wished to a larger degree to score better than other fellow students on the quiz (68% vs. 38% and 43%) , and felt to a larger degree increased pulse doing the quiz (46% vs. 13% and 8%).

We also performed an analysis on how the use of quiz method affects the engagement and motivation for male vs. female students separately. Table II summarizes the descriptive statistics and the results of the Kruskal-Wallis test for the statements with most noticeable differences between the genders (complete results from the tests and the descriptive statistics can be found in Table X and **Error! Reference source not found.**Table XI in Appendices).

Table II Gender Comparison of Paper, Clicker and Kahoot!

Statement	Group	Disagree	Neutral	Agree	H	P
2. I wanted to answer quiz without help from others (female students)	Paper	27%	23%	50%	3.88	0.1437
	Clicker	17%	22%	62%		
	Kahoot!	9%	14%	77%		
2. I wanted to answer quiz without help from others (male students)	Paper	11%	15%	74%	0.43	0.8065
	Clicker	6%	15%	79%		
	Kahoot!	8%	13%	79%		
3. I did the quiz only because the teacher told me to (female students)	Paper	36%	27%	36%	18.76	<0.0001
	Clicker	60%	23%	17%		
	Kahoot!	75%	23%	2%		
3. I did the quiz only because the teacher told me to (male students)	Paper	70%	21%	9%	6.16	0.046
	Clicker	54%	33%	14%		
	Kahoot!	79%	13%	8%		
6. The quiz was boring and not engaging (female students)	Paper	58%	34%	8%	6.21	0.0448
	Clicker	63%	20%	17%		
	Kahoot!	86%	7%	7%		
6. The quiz was boring and not engaging (male students)	Paper	62%	34%	4%	3.47	0.1764
	Clicker	64%	30%	6%		
	Kahoot!	82%	18%	0%		
10. I wished to have a better score on the quiz than most other students in the class (female students)	Paper	28%	42%	30%	9.93	0.007
	Clicker	18%	37%	45%		
	Kahoot!	5%	41%	55%		
10. I wished to have a better score on the quiz than most other students in the class (male students)	Paper	17%	34%	49%	15.74	0.0004
	Clicker	15%	44%	42%		
	Kahoot!	3%	13%	84%		
11. I felt increased pulse when I answered questions in the quiz (female students)	Paper	73%	20%	7%	10.6	0.005
	Clicker	71%	17%	12%		
	Kahoot!	43%	25%	32%		
11. I felt increased pulse when I answered questions in the quiz (male students)	Paper	53%	25%	23%	28.53	<0.0001
	Clicker	63%	32%	5%		
	Kahoot!	24%	13%	63%		

The following trends of differences between genders were identified:

- The female students wish to answer quiz without help was more related to use of technology than for male students (statement 2).
- A higher percentage of male student doing the paper quiz disagreed in that they did the quiz only because the teacher told them to compared to female students (statement 3).
- The motivation and engagement of female students have a stronger correlation to the quiz method used compared to male students (statement 6).
- A larger percentage of male students compared to female students focused on being competitive when playing Kahoot! (statement 10).
- Male students reported to a larger degree that they felt increased pulse when playing Kahoot! compared to female students.

3.2 Results from the Paper Quiz

In this section we look at how gender and whether the students play games or not, affect their engagement and motivation when doing a paper quiz to review a subject in a lecture.

Table III shows the descriptive statistics and the results from the Mann-Whitney test for the lecture using Paper quiz comparing Female vs. Male students for k=2 with n_{female}=74 and n_{male}=53.

Table III Gender Comparison for the Paper Quiz Lecture

Statement	Group	Disagree	Neutral	Agree	U_A	Z	P
1.I concentrated on the quiz to get correct answer	Female	9%	32%	57%	2099.5	-0.67	0.2514
	Male	6%	30%	64%			
2. I wanted to answer quiz without help from others	Female	27%	23%	50%	2452.5	-2.4	**0.0082**
	Male	11%	15%	**74%**			
3. I did the quiz only because the teacher told me to	Female	36%	27%	36%	1209	3.67	**0.0001**
	Male	**70%**	21%	9%			
4. To complete the quiz gave me satisfaction	Female	**43%**	31%	26%	2437.5	-2.33	**0.0099**
	Male	23%	36%	**42%**			
5. I enjoyed quiz so much I want to know more about the topic	Female	**50%**	35%	15%	2612.5	-3.18	**0.0007**
	Male	19%	53%	28%			
6. The quiz was boring and not engaging	Female	58%	34%	8%	1850.5	0.54	0.2946
	Male	62%	34%	4%			
7. It gave me satisfaction to complete quiz in a satisfactory way	Female	20%	47%	32%	2348.5	-1.89	**0.0294**
	Male	15%	32%	**53%**			
8. I do not think this lecture was worth my time and effort	Female	49%	31%	20%	1561	1.95	**0.0256**
	Male	**66%**	26%	8%			
9. To score high on the quiz was the most satisfactory in the lecture	Female	42%	39%	19%	2202.5	-1.18	0.119
	Male	32%	42%	26%			
10. I wished for better score on the quiz than most other students in class	Female	28%	42%	30%	2389.5	-2.09	**0.0183**
	Male	17%	34%	**49%**			
11. I felt increased pulse when I answered questions in the quiz	Female	73%	20%	7%	2413.5	-2.21	**0.0136**
	Male	53%	25%	**23%**			

Table III shows statistically significant differences between the genders ($p \leq 0.05$) for the statements 2, 3, 4, 5, 7, 8, 10, and 11. The general picture is that female students are less motivated and engaged by the paper quiz compared to male students. E.g., 70% of male students disagreed in the statement "I did the quiz only because the teacher told me to" vs. 36% for female students. Similarly, 53% of male students agree to the statement "it gave me satisfaction to complete the quiz in a satisfactory way vs. 32% for female students.

The questionnaire also included a question regarding the student played games or not (at least 1 hour a week). Table IV shows the descriptive statistics and the results from the Mann-Whitney test for lecture using a paper quiz comparing non-gaming vs. gaming students (more than 1 hour a week) for k=2 with $n_{non\text{-}gaming}=63$ and $n_{gaming}=64$. The gender distribution of non-gaming students were 78% female and 22% male, and for gaming students 39% female and 61% male.

Table IV Gaming Comparison for the Paper Lecture

Statement	Group	Disagree	Neutral	Agree	U_A	Z	P
1.I concentrated on the quiz to get correct answer	Non-G	8%	41%	51%	2379.5	-1.75	**0.0401**
	Gaming	8%	22%	**70%**			
2. I wanted to answer quiz without help from others	Non-G	30%	22%	48%	2560.5	-2.62	**0.0044**
	Gaming	11%	17%	**72%**			
3. I did the quiz only because the teacher told me to	Non-G	46%	21%	33%	1724	1.41	0.0793
	Gaming	55%	28%	17%			
4. To complete the quiz gave me satisfaction	Non-G	48%	35%	17%	2755	-3.56	**0.0002**
	Gaming	22%	31%	**47%**			
5. I enjoyed quiz so much I want to know more about the topic	Non-G	44%	38%	17%	2328.5	-1.5	0.0668
	Gaming	30%	47%	23%			
6. The quiz was boring and not engaging	Non-G	54%	35%	11%	1718.5	1.43	0.0764
	Gaming	66%	33%	2%			
7. It gave me satisfaction to complete quiz in a satisfactory way	Non-G	27%	43%	30%	2579	-2.71	**0.0034**
	Gaming	9%	39%	**52%**			
8. I do not think this lecture was worth my time and effort	Non-G	52%	25%	22%	1768	1.19	0.117
	Gaming	59%	33%	8%			
9. To score high on the quiz was the most satisfactory in the lecture	Non-G	43%	33%	24%	2130	-0.55	0.2912
	Gaming	33%	47%	20%			
10. I wished for better score on the quiz than most other students in class	Non-G	29%	40%	32%	2324	-1.48	0.0694
	Gaming	19%	38%	44%			
11. I felt increased pulse when I answered questions in the quiz	Non-G	**73%**	17%	10%	2363.5	-1.67	**0.0475**
	Gaming	56%	27%	17%			

Table IV shows statistically significant differences ($p \leq 0.05$) for five statements in the questionnaire. The results show that gaming students to a larger degree concentrate on the quiz to get a correct answer (statement 1), answer the quiz without help form others (statement 2), get more satisfied by completing the quiz (statement 4), are more satisfied from completing the quiz in a satisfactory way (statement 7), and feel increase pulse when answering questions in the quiz.

3.3 Results from the Clicker Quiz

In this section we look at how gender and whether the students play games or not affect their engagement and motivation when doing a Clicker quiz to review a subject in a lecture.

Table V shows the descriptive statistics and the results from the Mann-Whitney test for Clicker Quiz comparing Female vs. Male students for k=2 with $n_{female}=65$ and $n_{male}=110$.

Table V Gender Comparison of the Clicker Quiz Lecture

Statement	Group	Disagree	Neutral	Agree	U_A	Z	P
1.I concentrated on the quiz to get correct answer	Female	6%	17%	77%	3272.5	0.93	0.1762
	Male	7%	25%	68%			
2. I wanted to answer quiz without help from others	Female	17%	22%	62%	4241.5	-2.06	**0.0197**
	Male	6%	15%	**79%**			
3. I did the quiz only because the teacher told me to	Female	60%	23%	17%	3717	-0.44	0.33
	Male	54%	33%	14%			

4. To complete the quiz gave me satisfaction	Female	26%	48%	26%	3695	-0.37	0.3557
	Male	29%	48%	34%			
5. I enjoyed quiz so much I want to know more about the topic	Female	**43%**	35%	22%	4139.5	-1.74	**0.0409**
	Male	29%	41%	30%			
6. The quiz was boring and not engaging	Female	63%	20%	17%	3419	0.48	0.3156
	Male	64%	30%	6%			
7. It gave me satisfaction to complete quiz in satisfactory way	Female	18%	37%	45%	3624	-0.15	0.4404
	Male	15%	42%	44%			
8 I do not think this lecture was worth my time and effort	Female	51%	28%	22%	3448	0.39	0.3483
	Male	51%	35%	15%			
9. To score high on the quiz was the most satisfactory in the lecture	Female	34%	35%	31%	3387	0.58	0.281
	Male	36%	38%	25%			
10. I wished for better score on the quiz than most other students in class	Female	18%	37%	45%	3571	0.01	0.496
	Male	15%	43%	42%			
11. I felt increased pulse when I answered questions in the quiz	Female	71%	17%	12%	3755.5	-0.56	0.2877
	Male	63%	32%	5%			

Table V shows statistically significant differences ($p \leq 0.05$) only for two statements in the questionnaire. Male students agreed to a larger degree that they wanted to answer the quiz without the help of others and female students disagreed to a larger extent that they enjoyed the quiz so much they wanted to know more about the topic.

Table VI shows the descriptive statistics and the results of the Mann-Whitney test for lecture using Paper quiz comparing non-gaming vs. gaming students (more than 1 hour a week) for k=2 with $n_{non\text{-}gaming}$=67 and n_{gaming}=108. The non-gaming group consisted of 55% female and 45% male students, while the gaming group consisted of 26% female and 74% male students. None of the statements showed a statistically significant difference ($p \leq 0.05$) between non-gaming and gaming students. The statement with most difference was statement 11, where non-gaming students to a larger extent disagreed that they felt an increased pulse when doing the Clicker quiz.

Table VI Gaming Comparison of the Clicker Quiz Lecture

Statement	Group	Disagree	Neutral	Agree	U_A	Z	P_1
1.I concentrated on the quiz to get correct answer	Non-G	3%	21%	76%	3297	0.98	0.1635
	Gaming	9%	22%	69%			
2. I wanted to answer quiz without help from others	Non-G	9%	18%	73%	3567	0.16	0.4364
	Gaming	11%	17%	72%			
3. I did the quiz only because the teacher told me to	Non-G	57%	28%	15%	3652	-0.1	0.4602
	Gaming	56%	30%	15%			
4. To complete the quiz gave me satisfaction	Non-G	31%	36%	33%	3670	-0.16	0.4364
	Gaming	26%	44%	30%			
5. I enjoyed quiz so much I want to know more about the topic	Non-G	40%	36%	24%	3990.5	-1.14	0.1271
	Gaming	31%	41%	29%			
6. The quiz was boring and not engaging	Non-G	66%	21%	13%	3668.5	-0.15	0.4404
	Gaming	62%	30%	8%			
7. It gave me satisfaction to complete quiz in satisfactory way	Non-G	21%	30%	49%	3520	0.3	0.3821
	Gaming	13%	46%	41%			
8. I do not think this lecture was worth my time and effort	Non-G	57%	28%	15%	3966.5	-1.07	0.1423
	Gaming	47%	34%	19%			
9. To score high on the quiz was the most satisfactory in the lecture	Non-G	34%	36%	30%	3479.5	0.42	0.3372
	Gaming	36%	38%	26%			
10. I wished for better score on the quiz than most other students in class	Non-G	21%	40%	39%	3981.5	-1.11	0.1335
	Gaming	13%	41%	46%			
11. I felt increased pulse when I answered questions in the quiz	Non-G	75%	16%	9%	4079.5	-1.42	0.0778
	Gaming	60%	32%	7%			

3.4 Results from the Kahoot! Quiz

In this section we look at how gender and whether the students play games or not affects their engagement and motivation when doing a Kahoot! quiz to review a subject in a lecture. Table VII shows the descriptive statistics and the results from the Mann-Whitney test for Kahoot! Quiz comparing Female vs. Male students for k=2 with n_{female}=65 and n_{male}=110.

Table VII Gender Comparison of the Kahoot! Quiz Lecture

Statement	Group	Disagree	Neutral	Agree	U_A	Z	P
1.I concentrated on the quiz to get correct answer	Female	9%	14%	77%	861	-0.23	0.409
	Male	3%	18%	79%			
2. I wanted to answer quiz without help from others	Female	9%	25%	66%	938.5	-0.95	0.1711
	Male	8%	13%	79%			
3. I did the quiz only because the teacher told me to	Female	75%	23%	2%	815.5	0.19	0.4247
	Male	79%	13%	8%			
4. To complete the quiz gave me satisfaction	Female	27%	41%	32%	1012	-1.63	*0.0516*
	Male	16%	34%	*50%*			
5. I enjoyed quiz so much I want to know more about the topic	Female	36%	52%	11%	903.5	-0.62	0.2676
	Male	34%	45%	21%			
6. The quiz was boring and not engaging	Female	86%	7%	7%	865.5	-0.27	0.3936
	Male	82%	18%	0%			
7. It gave me satisfaction to complete quiz in satisfactory way	Female	14%	41%	45%	916	-0.74	0.2297
	Male	16%	26%	58%			
8. I do not think this lecture was worth my time and effort	Female	57%	30%	14%	722	1.06	0.1446
	Male	71%	18%	11%			
9. To score high on the quiz was the most satisfactory in the lecture	Female	20%	32%	48%	889.5	-0.49	0.3121
	Male	24%	18%	58%			
10. I wished for better score on the quiz than most other students in class	Female	5%	41%	55%	1080	-2.26	**0.0119**
	Male	3%	13%	**84%**			
11. I felt increased pulse when I answered questions in the quiz	Female	43%	25%	32%	1096	-2.41	**0.008**
	Male	24%	13%	**63%**			

Table VII shows statistically significant differences ($p \leq 0.05$) only for two statements in the questionnaire. Male students agreed more to be competitive (statement 10) compared to female students, and they also to a larger degree felt increase pulse when playing Kahoot! (statement 11). There is also a borderline statement where male students to a larger extent got satisfaction for completing the quiz (statement 4, p=0.0516).

Table VIII shows the descriptive statistics and the results of the Mann-Whitney test for lecture using Paper quiz comparing Non-gaming vs. Gaming students (more than 1 hour a week) for k=2 with $n_{non\text{-}gaming}$=41 and n_{gaming}=41. The non-gaming group consisted of 71% female and 29% male students, while the gaming group consisted of 37% female and 63% male students. The only statement with statistically significant difference ($p \leq 0.05$) is statement 10, where gaming students were more competitive than non-gaming students. There is also a borderline case for statement 3 where gaming students to a larger extend disagreed that they did the quiz only because the teacher told them to (p=0.0526).

Table VIII Gaming Comparison of the Kahoot! Quiz Lecture

Statement	Group	Disagree	Neutral	Agree	U_A	Z	P
1.I concentrated on the quiz to get correct answer	Non-G	12%	5%	83%	786	0.5	0.3085
	Gaming	0%	27%	73%			
2. I wanted to answer quiz without help from others	Non-G	7%	27%	66%	928.5	-0.81	0.209
	Gaming	10%	12%	78%			
3. I did the quiz only because the teacher told me to	Non-G	66%	29%	5%	665	1.62	*0.0526*
	Gaming	*88%*	7%	5%			
4. To complete the quiz gave me satisfaction	Non-G	24%	39%	37%	913	-067	0.2514
	Gaming	20%	37%	44%			
5. I enjoyed quiz so much I want to know more about the topic	Non-G	37%	46%	17%	844.5	-0.03	0.4880
	Gaming	34%	51%	15%			
6. The quiz was boring and not engaging	Non-G	83%	10%	7%	811	0.27	0.3936
	Gaming	85%	15%	0%			
7. It gave me satisfaction to complete quiz in satisfactory way	Non-G	17%	34%	49%	895.5	-0.51	0.3050
	Gaming	12%	34%	54%			
8. I do not think this lecture was worth my time and effort	Non-G	61%	29%	10%	819.5	0.19	0.4247
	Gaming	66%	20%	15%			
9. To score high on the quiz was the most satisfactory in the lecture	Non-G	24%	24%	51%	876	-0.32	0.3745
	Gaming	20%	27%	54%			
10. I wished for better score on the quiz than most other students in class	Non-G	5%	44%	51%	1123.5	-2.62	**0.0044**
	Gaming	2%	12%	**85%**			
11. I felt increased pulse when I answered questions in the quiz	Non-G	39%	17%	44%	911.5	-0.65	0.2578
	Gaming	29%	22%	49%			

3.5 Learning Outcome

We were able only to compare the learning outcome from the two parallel lectures where Paper and Kahoot! were used. The reason for this was that the results from the post-test in the Clicker quiz lecture were incomplete due to lack of time at the end of the lecture.

The Mann-Whitney test was used to investigate if there were any statistically significant difference between the performances of students using the Paper vs. students using Kahoot!. Results from the Mann-Whiney test along with the descriptive statistics are shown in Table IX. The Min, Max, Mean and Median shows improvement from pre-test to post-test in number correct answers in the test. There is a tendency for a higher mean value for the lecture with Kahoot! quiz compared to the lecture with the Paper quiz, but the difference is not statistically significant.

Table IX Learning Outcome from Paper Quiz vs. Kahoot! Quiz

Treatment	N	Min	Max	Mean	Median	U_A	Z	P
Paper	127	0	7	3.669	4	5655.5	-1.05	0.1469
Kahoot!	82	1	6	3.817	4			

4. Discussions

This section discusses the results presented in previous section and discusses some threats of validity.

4.1 Discussion of the Results

The results presented in Section 3 showed that there were some significant differences in terms of engagement and motivation for the type of quiz method used in a lecture. If we only

compare the usage of paper quizzes and the clicker system, no major differences were found in the motivation or engagement for the whole group of students. This is an interesting result, as it indicates that the students do not automatically get more motivated and engaged by introducing a digital student response system instead of a paper-based quiz. However, it should be noted that a student response system gives other benefits both for the teacher and the students. The teacher can benefit from collecting data and can respond directly to questions were the students lack knowledge when the questions are asked. The students can benefit from being able to give anonymous responses and the opportunity to get immediate and personal feedback on whether they have answered correctly or not. It is interesting to see the noticeable differences between a game-based student-response system and the two other methods for running reviews. The students that used Kahoot! for reviewing, were more engaged and motivated in general, actually enjoyed the quiz itself more, wanted to do better on the quiz, were more competitive, and felt to a larger degree increased pulse.

If we consider how motivation and engagement is related to gender for the three different methods of quizzing, no major differences were found. However, there is a tendency of more competitiveness among male students. If we investigate differences between the genders by looking at one quizzing method at a time, it is striking that the most noticeable difference between genders can be found for the paper quiz (see

Table III). Female students are significantly more negative to paper quizzes than male students. In over 70% of the statements there was a statistically significant difference, where female students answered more negative than male students. Both for the Clicker and the Kahoot! tools, only 18% of the statements had a statistically significantly difference for gender (Table V and Table VII). Based on these results, it seams like the use of a student response system reduces gender differences in engagement and motivation.

Regarding students that play video games vs. students that do not, there is a similar tendency. There are more differences found for the paper quiz compared to Clicker and Kahoot!. For the group that did the paper quiz, gaming students were more engaged and motivated compared to non-gaming students. We did not find this effect for the Clicker group (

Table VI) or the Kahoot! group (Table VIII). The only statistically significant difference spotted for the Kahoot! group of students, was that gaming students to a higher degree were more competitive than non-gaming students (Table VIII).

Regarding the learning outcome, the experiment did not produce any statistically significant difference between the paper quiz and Kahoot!. This was not a major surprise, as it is likely that differences in learning outcomes produced through higher engagement and motivation must been observed over some time, and not in a single lecture. In our study, we assigned 10 out of 45 minutes to the quiz. It would have been surprising if those 10 minutes had made a significant impact on the learning outcome.

4.2 Threats to Validity

This section addressed the most important threats to the validity of the results in the experiment described in this article.

4.2.1 Intern Validity

The intern validity of an experiment concerns "the validity of inferences about whether observed covariation between A (the presumed treatment) and B (the presumed outcome) reflects a causal relationship from A to B as those variables were manipulated or measured" (Shadish, Cook et al. 2002). If changes in B have causes other than the manipulation of A, there is a threat to the internal validity.

There are two main internal validity threats to this experiment. The *first internal threat* is that the sample of the three groups (Paper, Clicker and Kahoot!) in the experiment was not randomized. The three different groups all consist of first year university students that are divided into parallels based on the educational program they choose. None of the programs for the three groups are computer science or software engineering programs. All the three groups have similar distribution of gender as well as whether the students play games or not. To our knowledge, there should not be any major differences between the three groups.

The *second internal threat* is that there were any differences in how the three parallel lectures were conducted. For all the three parallels, the same teacher was teaching using the same agenda, the same slides, and the same questionnaires. We identify two differences between the parallels. The most obvious was the method for performing the quiz, which was a part of the experiment. The second minor difference was that the teacher ran out of time for the Clicker lecture, resulting in not enough time for all the students to do the post-test. Due to uncertain data on learning analysis, the Clicker lecture was removed from the learning outcome analysis.

4.2.2 Construct Validity

Construct validity concerns the degree to which inferences are warranted, from (1) the observed persons, settings, and cause and effect operations included in a study to (2) the constructs that these instances might represent. The question, therefore, is whether the sampling particulars of a study can be defended as measures of general constructs (Shadish, Cook et al. 2002). The goal of this experiment was to investigate how the method of quiz in a lecture affects the students' motivation, engagement and learning outcome. Our approach was to carry out a pre- and post-test to measure learning outcome, and do a questionnaire constructed on existing frameworks to investigate the students' engagement and motivation. Apart from the lack of data from the Clicker lecture on learning outcome, our approach for the experiment should be sufficient to give us the answers we were looking for.

4.2.3 External Validity

The issue of external validity concerns whether a causal relationship holds (1) for variations in persons, settings, treatments, and outcomes that were in the experiment and (2) for persons, settings, treatments, and outcomes that were not in the experiment (Shadish, Cook et al. 2002). Our experiment compares three different approaches for running quizzes in a lecture namely using paper forms, using a simple student response system (Clicker), and using a

game-based student response system (Kahoot!). The experiment was executed in an IT introductory course at a university in Norway. We do not believe that the results described in this article can be generalized to any use of quiz methods or for students of any age, but we believe that the results can be generalized if similar quiz approaches (paper forms, simple student response system and game-based student response system) are used in a university context.

5. Conclusions

In this article, we have presented an experiment to investigate how the usage of quizzes in review-lectures affects motivation, engagement and the learning outcome. Our experiment revealed that students were more engaged and motivated when a game-based student response system was used compared to when paper forms or a simple non-game-based student response system was used. Our study also showed that differences in motivation and engagement for female vs. male students depends on the technology used. Female students were more negative towards paper quizzes. The results also showed that the smallest differences between the genders were found when a game-based student response system was used. Similarly, the results from the experiment showed greater motivation and engagement differences between gaming and non-gaming students for paper quizzes than for the other two review methods. The study also revealed that male students and gaming students are more competitive than female students and non-gaming students respectively.

Regarding the learning outcome, the results from the experiment did not show any statistically significant differences between the quiz methods (only paper and Kahoot! was tested). In future studies, we will investigate more thoroughly whether the learning outcome varies by the method or by the quiz-tool used.

Acknowledgements

NTNU Discovery and the FORNY2020 program from Research Council of Norway have supported the research published in this article. The authors would like to thank the students from TDT4105 IT Introductory Class 2013. We would also like to thank Richard Taylor and Walt Scacchi at the Institute for Software Research (ISR) at the University of California, Irvine (UCI) for providing a stimulating research environment and for hosting a visiting researcher from Norway.

References

Anderson, J. and M. Barnett (2011). "Using video games to support pre-service elementary teachers learning of basic physics principles." Journal of Science Education and Technology **20**(4): 347-362.

Bessler, W. C. and J. J. Nisbet (1971). "The use of an electronic response system in teaching biology." Science Education **55**(3): 275-284.

Burnstein, R. A. and L. M. Lederman (2001). "Using wireless keypads in lecture classes." The Physics Teacher **39**(1): 8-11.

Caldwell, J. E. (2007). "Clickers in the large classroom: Current research and best-practice tips." CBE-Life Sciences Education **6**(1): 9-20.

Carnevale, D. (2005). "Run a class like a game show:'Clickers' keep students involved." Chronicle of Higher Education **51**(42): B3.

Carr, D. and T. Bossomaier (2011). "Relativity in a rock field: A study of physics learning with a computer game." Australasian Journal of Educational Technology **27**(6): 1042-1067.

Carver Jr, C. A., et al. (1999). "Enhancing student learning through hypermedia courseware and incorporation of student learning styles." Education, IEEE Transactions on **42**(1): 33-38.

Casanova, J. (1971). "An instructional experiment in organic chemistry. The use of a student response system." Journal of Chemical Education **48**(7): 453.

Coca, D. M. and J. Slisko (2013). "Software Socrative and Smartphones as Tools For Implementation of Basic Processes of Active Physics Learning in Classroom: An Initial Feasibility Study With Prospective Teachers." European Journal of Physics Education **4**(2).

Cutts, Q. I., et al. (2004). Maximising Dialogue in Lectures using Group Response Systems. CATE.

Gee, J. P. (2003). "What video games have to teach us about learning and literacy." Comput. Entertain. **1**(1): 20-20.

Gruenstein, A., et al. (2009). A self-transcribing speech corpus: collecting continuous speech with an online educational game. SLaTE Workshop.

Judson, E. (2002). "Learning from past and present: Electronic response systems in college lecture halls." Journal of Computers in Mathematics and Science Teaching **21**(2): 167-181.

Kay, R. H. and A. LeSage (2009). "Examining the benefits and challenges of using audience response systems: A review of the literature." Computers & Education **53**(3): 819-827.

Kebritchi, M., et al. (2010). "The effects of modern mathematics computer games on mathematics achievement and class motivation." Computers & Education **55**(2): 427-443.

Kirriemuir, J. and A. McFarlane (2004). "Literature Review in Games and Learning."

Lepper, M. R., et al. (2005). "Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates." Journal of educational psychology **97**(2): 184.

Liao, C. C., et al. (2011). "My‐Mini‐Pet: a handheld pet‐nurturing game to engage students in arithmetic practices." Journal of Computer Assisted Learning **27**(1): 76-89.

Lucas, A. (2009). "Using peer instruction and i-clickers to enhance student participation in calculus." Primus **19**(3): 219-231.

Malone, T. W. (1980). What Makes Things Fun to Learn? Heuristics for designing Instructional Computer Games. The 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems. Palo Alto, California, United States, ACM Press.

Miller, L., et al. (2002). "Teaching neuroscience through web adventures: adolescents reconstruct the history and science of opioids." The Neuroscientist **8**(1): 16-21.

Papastergiou, M. (2009). "Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation." Computers & Education **52**(1): 1-12.

Pintrich, P. R. (1991). "A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)."

Prensky, M. (2005). "Computer games and learning: Digital game-based learning." Handbook of computer game studies **18**: 97-122.

Rosas, R., et al. (2003). "Beyond Nintendo: design and assessment of educational video games for first and second grade students." Computer Education **40**(1): 71-94.

Schell, J., et al. (2013). "Catalyzing learner engagement using cutting-edge classroom response systems in higher education." Cutting-edge Technologies in Higher Education **6**: 233-261.

Sellar, M. (2011). "Poll everywhere." The Charleston Advisor **12**(3): 57-60.

Shadish, W. R., et al. (2002). Experimental and quasi-experimental designs for generalized causal inference, Wadsworth Cengage learning.

Sharples, M. (2000). "The design of personal mobile technologies for lifelong learning." Comput. Educ. **34**(3-4): 177-193.

Squire, K. (2005). "Changing the game: What happens when video games enter the classroom." Innovate: Journal of online education **1**(6).

Tüysüz, C. (2009). "Effect of the computer based game on pre-service teachers' achievement, attitudes, metacognition and motivation in chemistry." Sci Res Essays **4**(8): 780-790.

Wang, A. I., et al. (2007). Lecture Quiz - A Mobile Game Concept for Lectures. IASTED International Conference on Software Engineering and Application (SEA 2007). Cambridge, MA, USA, Acta Press**:** 6.

Wang, A. I., et al. (2008). An Evaluation of a Mobile Game Concept for Lectures. Proceedings of the 2008 21st Conference on Software Engineering Education and Training - Volume 00, IEEE Computer Society.

Wu, B., et al. (2011). IMPROVEMENT OF A LECTURE GAME CONCEPT-Implementing Lecture Quiz 2.0. Proceedings of the 3rd International Conference on Computer Supported Education.

Appendices

A1. Comparison of Quiz Methods for Female Students

Kruskal-Wallis test comparing Paper, Clicker and Kahoot! for k=3 with n_{paper}=74, $n_{clicker}$=65, $n_{Kahoot!}$=44, df=2 (Female students)

Table X Comparing the Use of Paper, Clicker and Kahoot! Female Students

Statement	Group	Disagree	Neutral	Agree	H	P
1.I concentrated on the quiz to get correct answer	Paper	9%	32%	58%	4.29	0.1171
	Clicker	6%	17%	77%		
	Kahoot!	9%	14%	77%		
2. I wanted to answer quiz without help from others	Paper	27%	23%	50%	3.88	0.1437
	Clicker	17%	22%	62%		
	Kahoot!	9%	14%	77%		
3. I did the quiz only because the teacher told me to	Paper	36%	27%	36%	18.76	<0.0001
	Clicker	60%	23%	17%		
	Kahoot!	75%	23%	2%		
4. To complete the quiz gave me satisfaction	Paper	43%	31%	26%	2.62	0.2698
	Clicker	26%	48%	26%		
	Kahoot!	27%	41%	32%		
5. I enjoyed the quiz so much that I want to know more about the topic	Paper	50%	35%	15%	1.18	0.5543
	Clicker	43%	35%	22%		
	Kahoot!	36%	52%	11%		
6. The quiz was boring and not engaging	Paper	58%	34%	8%	6.21	0.0448
	Clicker	63%	20%	17%		
	Kahoot!	86%	7%	7%		
7. It gave me satisfaction to complete the quiz in a satisfactory way	Paper	20%	47%	32%	2.16	0.3396
	Clicker	18%	37%	45%		
	Kahoot!	14%	41%	45%		
8. I do not think this lecture was worth my time and effort	Paper	49%	31%	20%	0.88	0.644
	Clicker	51%	28%	22%		
	Kahoot!	57%	30%	14%		
9. To score high on the quiz was the most satisfactory in the lecture	Paper	42%	39%	19%	9.55	0.0084
	Clicker	34%	35%	31%		
	Kahoot!	20%	32%	48%		
10. I wished to have a better score on the quiz than most other students in the class	Paper	28%	42%	30%	9.93	0.007
	Clicker	18%	37%	45%		
	Kahoot!	5%	41%	55%		
11. I felt increased pulse when I answered questions in the quiz	Paper	73%	20%	7%	10.6	0.005
	Clicker	71%	17%	12%		
	Kahoot!	43%	25%	32%		

A2. Comparison of Quiz Methods for Male Students

Kruskal-Wallis test comparing Paper, Clicker and Kahoot! for k=3 with n_{paper}=53, $n_{clicker}$=110, $n_{Kahoot!}$=38, df=2

Table XI Comparing the Use of Paper, Clicker and Kahoot! Male Students

Statement	Group	Disagree	Neutral	Agree	H	P
1.I concentrated on the quiz to get correct answer	Paper	6%	30%	64%	1.58	0.4538
	Clicker	7%	25%	68%		
	Kahoot!	3%	18%	79%		
2. I wanted to answer quiz without help from others	Paper	11%	15%	74%	0.43	0.8065
	Clicker	6%	15%	79%		
	Kahoot!	8%	13%	79%		
3. I did the quiz only because the teacher told me to	Paper	70%	21%	9%	6.16	0.046
	Clicker	54%	33%	14%		
	Kahoot!	79%	13%	8%		
4. To complete the quiz gave me satisfaction	Paper	23%	36%	42%	3.77	0.1518
	Clicker	29%	37%	34%		
	Kahoot!	16%	34%	50%		
5. I enjoyed the quiz so much that I want to know more about the topic	Paper	19%	53%	28%	1.77	0.4127
	Clicker	29%	41%	30%		
	Kahoot!	34%	45%	21%		
6. The quiz was boring and not engaging NBNB!!!	Paper	62%	34%	4%	3.47	0.1764
	Clicker	64%	30%	6%		
	Kahoot!	82%	18%	0%		
7. It gave me satisfaction to complete the quiz in a satisfactory way	Paper	15%	32%	53%	1.36	0.5066
	Clicker	15%	42%	44%		
	Kahoot!	16%	26%	58%		
8. I do not think this lecture was worth my time and effort	Paper	66%	26%	8%	4.64	0.0983
	Clicker	51%	35%	15%		
	Kahoot!	71%	18%	11%		
9. To score high on the quiz was the most satisfactory in the lecture	Paper	32%	42%	26%	8.02	0.0181
	Clicker	36%	38%	25%		
	Kahoot!	24%	18%	58%		
10. I wished to have a better score on the quiz than most other students in the class	Paper	17%	34%	49%	15.74	0.0004
	Clicker	15%	44%	42%		
	Kahoot!	3%	13%	84%		
11. I felt increased pulse when I answered questions in the quiz	Paper	53%	25%	23%	28.53	<0.0001
	Clicker	63%	32%	5%		
	Kahoot!	24%	13%	63%		