Reaction vs. Completeness in Game-based Learning - Comparing two game modes in a game-based student response system

Alf Inge Wang and Tintin Trong Hoang
Dept. of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway
alfw@idi.ntnu.no, tintin.hoang@outlook.com

Abstract: This paper describes the results of an experiment comparing two game modes in the game-based response system Quiz-a-tron. The goal of the experiment was to investigate how two different game modes affected the players’ engagement, perceived learning, concentration, motivation, and social experience. In the game mode Reaction, your score depends on how quickly you give one correct answer. In the game mode Completeness, your score depends on how complete you answer is with a score multiplier for giving correct answers in a row. In the experiment, 30 students were randomly chosen. The subjects first played Quiz-a-tron in one game mode, then filled out a questionnaire, and then played the game in the other game mode and filled out the questionnaire comparing both game modes. Half of the subjects started with the Reaction game mode, while the other half started with the Completeness game mode.

The results show that both game modes were perceived as fun, engaging, motivating, educational, stimulated to concentration, and provided a good social experience. Further, Completeness was perceived as more fun, motivating and a better social experience, and the students perceived that they learned more. The observations showed very different student behavior in the two game modes. The Reaction game mode was characterized by more students guessing answers, also for questions requiring mental calculations. In this game mode, most students gave their answer within the first 3 seconds when the time limit was 30 seconds. The “dead time” after giving answers was used to social bonding discussing questions and answers. The Completeness game mode made the students think more carefully about their answers and the pace of the game was slower. There was noticeably less frustration in this game mode. Also it was easier to fight you way back to the top of the scoreboard if the multipler was not lost. The observations also showed also the Completeness game mode produced less social interaction during the game.

Keywords: Game-based student response system, student engagement, game mode, evaluation.

1. Introduction

Student response systems (SRSs) have been around since the sixties (Judson 2002) and have been using teaching biology and chemistry since the early seventies (Bessler and Nisbet 1971, Casanova 1971). Initially SRSs were based on special hardware allowing 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 introduction of the Bring Your Own Device wave paved the way for a new generation of SRSs, where students can use their own devices to respond. Especially after the introduction of smart phones and tablets, easy access to wireless networks 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 (Sellars 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. One of the main aims of using SRSs is to boost student engagement, motivation, concentration and learning, which is also the goal of game-based learning (Gee 2005). The term game-based SRS (GSRS) was introduced through the release of the game-based learning platform Kahoot! (Wang 2015). The main difference between a GSRS and a SRS is that the game-based version focuses more on engaging and motivating the students through attractive graphical user-interfaces and audio, as well by gamifying the whole student response experience. In Kahoot! the gamification is done by temporarily transforming the classroom into a game show as shown on TV, where the teacher plays the role of a game show host and the students are the competitors. Kahoot! was designed from ground up as a game, but there are other SRSs which provide game elements as well. Socrative offers Space Race, which is a friendly competition where the goal is to get to be the first individual or group to fulfil a progress bar across the screen by answering correctly on questions (Méndez and Slisko 2013). Another example is TurningPoint,
which is a PowerPoint plug-in where the students give their responses using wireless clickers that includes scores and a scoreboard (Pettit, McCoy et al. 2015). Experiments with TurningPoint showed that a significant majority of the respondents agreed or strongly agreed that the games were engaging, and an effective learning tool. Similarly, Quizlet provides badges and achievements to motivate the user (Gruenstein, McGraw et al. 2009).

There are some examples on how SRSs with game elements affect the students’ attitude such as (Carnevale 2005, Martyn 2007, Wang, Øfsdal et al. 2007, Wang, Øfsdal et al. 2008, Wu, Wang et al. 2011, Pettit, McCoy et al. 2015, Wang 2015, Wang, Zhu et al. 2016). These papers describe studies where the system as a whole is evaluated and does not consider how individual game elements affect the SRS experience. There are few studies studying the effect of isolated game mechanisms in a learning context. Some studies have emerged testing the psychological and behavioral impacts of individual game design elements in non-game settings (Deterding, Dixon et al. 2011) using controlled experiments (Attali and Arieli-Attali 2015, Lieberoth 2015, Mekler, Brühlmann et al. 2015). For instance, Lieberoth studied the psychological effects of competitive game mechanics created by game-like look of an activity. However, as far as we know there is only one study on the isolated effects of various game elements in a game-based student response system. This study investigates how the use of points and audio affect the students experience in the class (Wang and Lieberoth 2016). The results from the experiment revealed that there are some significant differences whether audio and points are used in game-based learning in the areas of concentration, engagement, enjoyment, and motivation. The most surprising finding was how the classroom dynamics was positively affected by the use of audio.

This paper presents the results of an experiment that investigates how the students’ attitude is affected by the choice between two game modes with variations in how answers are given and how answers are awarded with points. The goal of this experiment is to optimize the game experience to maximize the students’ engagement, motivation, concentration, social experience and learning. As far as we know, this study is so far unique in its kind.

2. Quiz-a-tron

Quiz-a-tron is a simple game-based student response system (GSPS) similar to Kahoot!, Socrative and Quizlet, specifically developed to investigate how the way the students give their answers will affect their experience. The game is played in a classroom using the existing infrastructure such as a teacher’s laptop, a video projector, a wireless network, and various student devices (laptops, tablets and smartphones). The teacher client shows login information, the questions, and a scoreboard on a large screen the whole class can see. The students give their answers using any digital device with an Internet connection and a web-browser. The game play is similar to Kahoot!, where a question will be presented to the students, the students give their answer among several alternatives, and a scoreboard shows the students with highest scores between each question. At the end, a winner with the highest score will be announced.

Quiz-a-tron uses a web-based client-server architecture, and was implemented using Express web application framework, Node.js, Socet.IO, jQuery Mobile, ClassyCountdown, and FastClick.js. The reason we did not use any existing platforms such as Kahoot!, was to ensure to have full control over the application, and provide a game-based student response system without any addition features or artefacts that could affect the experimental results. Also none of the existing GSRSSs provide both the game modes Reaction and Completeness, which were the focus of our research. Figure 1 shows four screenshots from the teacher client: The game mode selection screen a), the start game screen b), the question screen c), and the scoreboard screen d).
Figure 1: Screenshots from the teacher client

Figure 2 shows four screenshots of the student client: the login screen a), the answer screen for the Reaction game mode b), the answer screen for the Completeness game mode c), and the score screen d). Note that unlike Kahoot! where both the questions and the answers are shown on the teacher client and not on the student client, in Quiz-a-tron, only the question is shown on the teacher client while the answers are only shown on the student client.

Figure 2: Screenshots from the student client

The remaining part of this section will describe the two game modes in Quiz-a-tron: Reaction and Completeness.

### 2.1 Reaction Game Mode

The Reaction game mode is the baseline game mode, which is the same game mode used in e.g. Kahoot!, where the goal is for the player to give a correct answer as quickly as possible to get a high score. This game mode focuses on reaction time of the players to create excitement and engagement. In this game mode, it is only possible to give one correct answer, even if there are several correct answers among the alternatives. The score is calculated by combining a base score for getting the correct answer, as well as a bonus score which is calculated by taking the base score and multiply it with the percentage of time left with two decimal precision.

The code for calculating the score is the following:
The base score was set to 500 points. This means if a player submits a correct answer after 3 seconds on a 10 second question, the player will be awarded with $500+500\times(1-\frac{3}{10}) = 850$ points.

When a player has submitted an answer, it will not be possible to change the answer.

### 2.2 Completeness Game Mode

The Completeness game mode encourages the player to give as complete answers as possible, and the score does not depend on the time used as long as the answer is given within the allocated timeframe. The players are presented with a question where at least one out of four answers is correct. The player can state that more than one answer alternative is correct, by choosing from one to four correct answer alternatives for one question. If none of the answers submitted are wrong, the player will be awarded a score of the base score multiplied with the percentage of how many correct answers she or he was able to identify. This means e.g. for a question with three correct answer alternatives, a player would get 33% score if she or he give one correct answer, 66% score if she or he gives two correct answers, and 100% if she or he gives three correct answers. In addition, giving correct answers in a row will give extra points through a multiplier which will increase over time. However, if any of the submitted answers are incorrect, the player will be awarded with 0 points, and the score bonus multiplier will be reset. This means that a player will get 0 points, if she or he answer one incorrect answer, or e.g. if she or he answer two correct answers and one incorrect answer. In a sense, one could say that the completeness game mode focuses on getting both correct and complete answers, as well as encouraging answering correctly on all questions. It is also possible not to give any answer, which will award the player with 0 points, but the score bonus multiplier will retain. The player is also free to change their choice of answers as many times as they like until the time runs out. The score will be computed when the answer time runs out.

The code for calculating the score for the completeness game mode is the following:

```java
if (wrongAnswers==0) {
    score += (correctAnswers/maxAnswers).toFixed(2) * questionBaseAward * multiplier;
    multiplier += 1;
} else {
    multiplier = 1;
}
```

If a player submits two correct answers out of three correct answers, and the base score is 500 and the player has answered the last question correctly thus having a multiplier of 2, the player will be awarded $\frac{2}{3}\times500\times2 = 670$ points.

The Completeness game mode experiments with the players’ willingness to take risks on submitting answers they are not completely sure about in hope of gaining more points. As there the score bonus multiplier is lost if an incorrect answer is given, our hypothesis is that this game mode should provide the same tension as the Reaction game mode does.

### 3. Experiment and Results

The research goal of the study described in this paper was to investigate how two different game modes in a game-based student response system (GSRS) affect the students’ engagement, perceived learning, concentration, motivation, and social experience. The research method used is based on the Goal, Question Metrics (GQM) approach (Basili 1992) where we first define a research goal (conceptual level), then define a set of research questions (operational level), and finally describe a set of metrics to answer the research questions (quantitative level). In our case, the metrics used to give answers to the research questions are a mixture of quantitative (questionnaire) and qualitative data (observation and interview).

The research goal of this study was defined as the following using the GQL template:
The purpose of this study was to evaluate the effect of using two different game modes in a game-based student response system from the point of view of a student in the context of a lecture.

The following research questions (RQs) were defined by decomposing the research goal:
- RQ1: How is engagement affected by choice of game mode in a GSRS.
- RQ2: How is perceived learning affected by choice of game mode in a GSRS.
- RQ3: How is concentration affected by choice of game mode in a GSRS.
- RQ4: How is motivation affected by choice of game mode in a GSRS.
- RQ5: How is the social experience affected by choice of game mode in a GSRS.

The choice of game mode here is divided into one focusing on reaction and giving a single answer, and one focusing on giving as complete answer as possible and answering correct over time.

3.1 Testing Procedure and Experimental Design

Thirty test subjects were randomly chosen from engineering students at the Norwegian University of Science and Technology (NTNU). The sample of test subjects was appropriate as the Quiz-a-tron tool was designed to be game-based learning tool design for university students to review knowledge before a test or exam. The average age of the subjects was 22.7 years old and the gender distribution was 27% female and 73% male. Further 83% of the subjects said they play video games often in their spare time. On the question about what type of gamers they were, 33% said they focus on competition, 30% said they focus on collaboration, 20% said they focus on exploring, and 17% said they focus on socialization (Bartle 1996). The questions used in the quiz were based on the first year curriculum mandatory for all engineering students at NTNU. The experiment was carried out in a classroom at the university.

Figure 3 illustrates the design of the experiment. First, the 30 subjects were assigned into two groups of same size. Both groups got an introduction to the experiment and the Quiz-a-tron game. In the first phase, group 1 played Quiz-a-tron in the Reaction game mode, while group 2 played the Completeness game mode. Both groups were observed during this phase. After finishing the game, both groups answered the same questionnaire on enjoyment, engagement, perceived learning, concentration, motivation, emotional involvement, and social experience. Next, group 1 played Quiz-a-tron in the Completeness game mode, while group 2 played in the Reaction game mode. Both groups were also observed during this phase. After finished playing, both groups answered a comparison questionnaire where they should indicate which game mode they thought was most fun, engaging, motivating, emotional involving, gave best learning, requested most concentration, and provided the best social experience. Our experimental design makes sure that both groups get the same amount of time on each game mode and take into account the effect of trying the game for the first time as the subjects are divided into two groups of same size starting out with two different game modes. In addition to the design process shown in Figure 3, the subjects were asked to fill out a System Usability Scale (SUS) form (Brooke 1996) to make sure that there were no major usability issues that could ruin the experimental results. Quiz-a-tron got a SUS score of 81, which is a high usability score, which indicates that there were no major usability issues.
3.2 Results from the game mode-specific Questionnaire

After playing the first quiz in Quiz-a-tron, the subjects were asked to fill out a questionnaire on how they perceived the game-based student response system. Half of the subjects played the Reaction game mode, where the rest played the Completeness game mode. Table 1 shows the summarized results from this questionnaire. To enhance readability in the table, the 5-level Likerts scale (Likert 1932) from Strongly disagree to Strongly agree was grouped into Disagree (includes Strongly disagree and Disagree), Neutral, and Agree (Agree and Strongly agree).

Table 1: Results from game mode-specific questionnaire

<table>
<thead>
<tr>
<th>#</th>
<th>Statement</th>
<th>Game mode</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am having fun playing the game mode</td>
<td>Reaction</td>
<td>0%</td>
<td>7%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>0%</td>
<td>13%</td>
<td>87%</td>
</tr>
<tr>
<td>2</td>
<td>I am engaged when playing the game mode</td>
<td>Reaction</td>
<td>0%</td>
<td>7%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>0%</td>
<td>7%</td>
<td>93%</td>
</tr>
<tr>
<td>3</td>
<td>I am learning when playing the game mode</td>
<td>Reaction</td>
<td>0%</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>0%</td>
<td>13%</td>
<td>87%</td>
</tr>
<tr>
<td>4</td>
<td>I am concentrating when playing the game mode</td>
<td>Reaction</td>
<td>0%</td>
<td>7%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>0%</td>
<td>13%</td>
<td>87%</td>
</tr>
<tr>
<td>5</td>
<td>I am motivated to learn when playing the game mode</td>
<td>Reaction</td>
<td>0%</td>
<td>47%</td>
<td>53%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>7%</td>
<td>33%</td>
<td>60%</td>
</tr>
<tr>
<td>6</td>
<td>I am emotionally involved when playing the game mode</td>
<td>Reaction</td>
<td>0%</td>
<td>7%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>0%</td>
<td>40%</td>
<td>60%</td>
</tr>
<tr>
<td>7</td>
<td>I have a good social experience playing the game mode</td>
<td>Reaction</td>
<td>27%</td>
<td>7%</td>
<td>67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>7%</td>
<td>20%</td>
<td>73%</td>
</tr>
</tbody>
</table>

Table 1 shows that there were no major differences in how the subjects perceived the two game modes with little variation between the two for most of the statements. In general, the large majority of the subjects perceived both game modes to be fun, engaging, that they learned from playing, that they concentrated when playing, and that it was a good social experience. The results were less positive related to if they were motivated to learn when playing the game. The only statement where there was a noticeable difference was for statement 6, where 93% playing the Reaction game mode said that they were emotionally involved when playing the game mode compared to 60% for the Completeness game mode. Our assumed hypothesis was that there should not be perceived difference in how the students perceived the two game modes. Thus, we ran a Mann-Whitney test to see if there were any statistically significant differences. For all the statements apart from statement 6, the P values were P>0.30. The Mann-Whitney test for statement 6 on emotional involvement when playing the game mode for N_A = N_B = 15 gave P=0.063 (U_A = 75, Z=1.53).

3.3 Results from comparing the two game modes

After the subjects had played Quiz-a-tron using both game modes, they were asked to fill out a questionnaire to state which game mode they preferred for different statements. Table 2 shows the results from the questionnaire comparing the two game modes.

Table 2: Results comparing the game modes

<table>
<thead>
<tr>
<th>#</th>
<th>Statement</th>
<th>Reaction</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>I had most fun playing...</td>
<td>33%</td>
<td>67%</td>
</tr>
<tr>
<td>9</td>
<td>I was most engaged playing...</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>10</td>
<td>I learned the most when playing...</td>
<td>17%</td>
<td>83%</td>
</tr>
<tr>
<td>11</td>
<td>I was most concentrated when playing...</td>
<td>40%</td>
<td>60%</td>
</tr>
<tr>
<td>12</td>
<td>I was most motivated when playing...</td>
<td>23%</td>
<td>77%</td>
</tr>
<tr>
<td>13</td>
<td>I was most emotionally involved playing...</td>
<td>47%</td>
<td>53%</td>
</tr>
<tr>
<td>14</td>
<td>I had the best social experience playing...</td>
<td>17%</td>
<td>83%</td>
</tr>
</tbody>
</table>

The results show that the majority of the subjects stated that the Completeness game mode was most fun, motivating, gave the best social experience, and they learned the most from this game mode. We also
calculated binomial probabilities to calculate the P for the sample size of 30 when our assumption was the two games would be perceived to have similar characteristics for the students. Our binomial probabilities calculations calculated the probability (P-value) for k or more out of 30 subjects would prefer one game mode over the other (n=50, p=0.5, q=0.5, mean=15, var=7.5, stdev=2.7386). Table 3 shows the binomial probabilities for choosing the Completeness game mode for the statements 8 to 14. The results show a P-value less than 0.05 for the statements 8, 10, 12, and 14.

Table 3: Binomial probabilities of for choosing the Completeness game mode for statements 8-14

<table>
<thead>
<tr>
<th>Nr</th>
<th>Statement</th>
<th>n</th>
<th>k</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>I had most fun playing...</td>
<td>30</td>
<td>20</td>
<td>0.04936</td>
</tr>
<tr>
<td>9</td>
<td>I was most engaged playing...</td>
<td>30</td>
<td>15</td>
<td>0.57223</td>
</tr>
<tr>
<td>10</td>
<td>I learned the most when playing...</td>
<td>30</td>
<td>25</td>
<td>0.00016</td>
</tr>
<tr>
<td>11</td>
<td>I was most concentrated when playing</td>
<td>30</td>
<td>18</td>
<td>0.18079</td>
</tr>
<tr>
<td>12</td>
<td>I was most motivated when playing...</td>
<td>30</td>
<td>23</td>
<td>0.00261</td>
</tr>
<tr>
<td>13</td>
<td>I was most emotionally involved playing...</td>
<td>30</td>
<td>16</td>
<td>0.42777</td>
</tr>
<tr>
<td>14</td>
<td>I had the best social experience playing...</td>
<td>30</td>
<td>25</td>
<td>0.00016</td>
</tr>
</tbody>
</table>

3.4 Observations

This section describes the observations and feedback from students playing the two game modes.

3.4.1 Observations from playing the Reaction game mode

One observation made for the Reaction game mode was that most players submitted their answer within the first 3 seconds when the time limit for giving an answer was set to be 30 seconds. As the computation of score in the Reaction game mode is only about being quick (as well as correct), the top priority of most students is to just read through the question and answers as fast as possible and picking the first answer that seems to be correct without any deep evaluation. This caused often to frustration when they realized they had hastily picked an incorrect answer. The reason for this behavior is that if you spend too much time to think about the answers, you will get a much lower score and it would be impossible to reach the top of the scoreboard. When players were given questions that required mental calculations, most just picked a random answer without giving any effort into solving it.

Another problem discovered in the Reaction game mode was that players who answered quickly experienced “dead time” where they waited for the others to finish. This affected their immersion and learning, and many said that they did not really think about the questions after answering since it was not possible to undo your submission. The only learning took place when the correct answer was revealed, and the discussions related to the correct answer among students.

Interestingly, the “dead time” was helpful in promoting social bonding. When students knew they had answered incorrectly, they often had an outburst of frustration and laughed together with other fellow students. This helped increase the engagement in the game.

The students also said that this game mode was a bit superficial and shallow. It was hard to be motivated to learn when it was challenging to invest time considering the questions due to how the score was computed. For most students, you either knew the correct answer of the top of your head, or you guessed the first that seemed to be correct.

3.4.2 Observations from playing the Completeness game mode

The Completeness game mode gave a different student behavior compared to the Reaction game mode. This game mode has a strong emphasis punishing wrong answers due to the bonus multiplier. In this game mode, the students had to be careful with the answers, since one incorrect answer could jeopardize the correct answers you gave for the same question (giving multiple answers).

The frustration level was much lower when students were playing the Completeness game mode. Players felt that it was easier to show their skills in the subject instead of being tested on how fast they could read and
respond. The motivation was also higher since players that did not know much about the subject could stick to the answers that were obvious, while others were able to show off their knowledge by correctly identifying all the correct answers.

Another feedback received from playing the Completeness game mode was that they felt that this game mode was more engaging since the comeback potential was higher. Despite having few points early in the game, a player could easily climb the scoreboard if they did not lose their bonus multiplier.

One distinct observation during the Completeness game mode was that the students were less social during the questions. Since you could change answers until the time ran out, the most competitive players had little incentive to discuss the questions with their neighbors.

3.4.3 Comparison
If we compare the two game modes, an analogy could be that Reaction compares to doing a sprint, while Completeness compares to doing a marathon. In Reaction, the focus is on answering one particular question as fast as possible, while in Completeness, you have to plan for the long run. Many of the students testing both game modes enjoyed the slower pace of Completeness, as they got more time to recall what they have learned and could focus more on getting the correct answers.

4. Conclusion
In this paper we have investigated how the two game modes Reaction and Completeness affects the students’ attitude through an experiment where 30 students tested the GSRS Quiz-a-tron. Overall, the students positively received both game modes. In terms of engagement (RQ1), there were no major differences between the two game modes apart from that the students perceived the Completeness game mode to be more fun. Related to perceived learning (RQ2), the students stated that they learned the most from the Completeness game mode. This result was backed up by observations and feedback from the students, as they spent more concentrated time thinking through the question before giving an answer. Especially the bonus multiplier made the students giving their answers more carefully in the Completeness game mode compared to the tendency to just pick the first answer that comes to mind in the Reaction game mode. In terms of the students’ concentration (RQ3), no significant differences were found. However, for motivation (RQ4), the students stated that they were more motivated playing the Completeness compared to the Reaction game mode. One possible reason for this was that the Completeness game mode made it easier to climb the scoreboard being persistent and determined. Finally, the experiment showed that the students preferred the Completeness to the Reaction game mode in terms of the social experience (RQ5). The Reaction game mode produced more audible interaction in the classroom, but a more stressful experience than playing the Completeness game mode.

Although our results showed that the Completeness game mode performed better than the Reaction game mode, we believe that variation in a GSRS is the way to go. By providing several game modes, it is possible to play through the same quizzes providing different experiences. As repetition is good for learning, having multiple game modes resulting in different students’ experiences is a good thing.

References


Likert, R. (1932). "A technique for the measurement of attitudes." *Archives of psychology*.


