Impact of Game-Based Student Response Systems on Factors of Learning in a Person-Centered Flipped Classroom on C Programming

Dominik Dolezal
Department of Computer Science, University of Applied Sciences Technikum Wien, Austria
IT Department, TGM – Vienna Institute of Technology, Austria
dominik.dolezal@technikum-wien.at

Alexandra Posekany
Department for Clinical Neurosciences and Preventive Medicine, Danube University Krems, Austria
IT Department, TGM – Vienna Institute of Technology, Austria

Renate Motschnig
Faculty of Computer Science, University of Vienna, Austria

Theresa Kirchweger
Department of Information Engineering and Security, University of Applied Sciences Technikum Wien, Austria

Robert Pucher
Department of Computer Science, University of Applied Sciences Technikum Wien, Austria

Abstract: Student response systems allow learners to actively contribute to the in-class learning process, while the teacher gets a picture of students' understanding of the topics at hand. However, not only the teacher is given an evaluation tool, but also the students are able to evaluate their own progress, which is an important aspect in person-centered learning. This paper analyzes the impact of the game-based student response system Kahoot! on a person-centered flipped university course on the introduction to programming and gives recommendations on using quizzes in class. The course was evaluated via anonymous questionnaires before and after introducing regular gamified quizzes. The students' feedback was extraordinarily positive; the results showed a significant improvement in the satisfaction with the course structure, course design, and perceived suitability of the course material. Hence, evidence found in this research suggests that game-based student response systems provide additional value to flipped classrooms in engineering education.

Introduction

Flipping the classroom and gamification are popular concepts in engineering education. In a flipped classroom, activities which are traditionally done in the classroom are instead done at home, while activities typically done as homework are performed in class (Bergmann & Sams, 2012, pp. 13-15, 62). Due to the asynchronous instruction, students can work at home at their own pace, allowing to answer questions, practice, and do lab work during the face-to-face lessons. Kanelopoulos et al. conducted a literature review on the topic of flipped classrooms in engineering education, stating that "flipping the classroom in engineering classrooms is a proposal that has taken a lot of attention lately" (2017, p. 19), allowing in-class activities to focus on higher-level thinking processes. They flipped two lessons of their machine design course and found evidence of flipped classrooms "promoting students' active involvement in the classroom and increase of their self-confidence" (Kanelopoulos et al., 2017, p. 32).

The term gamification can be defined in multiple ways. Two common definitions are the following: "process of game-thinking and game mechanics to engage users and solve problems" (Zichermann & Cunningham, 2011, p. xiv) as well as "application of game design principles and mechanics to non-game environments" (Kumar & Herger, 2017, p. 8). According to P&S Market Research, the global gamification market is expected to grow at a compound annual growth rate of 41.8%, reaching $22.9 billion by 2022 (P&S Market Research, 2016). MarketsandMarkets Research drew a similar conclusion and predicted an annual growth rate of 46.3% for the gamification market with an estimate of $11.10 billion by 2020 (MarketsandMarkets Research, 2016). The educational sector is expected to be affected in particular: Research and Markets forecast the education gamification market to grow at an annual growth rate of 66.2% (Research and Markets, 2016), predicting that it will reach $1.2 billion by 2020 (Banville, 2016).

Due to this notable trend towards gamification in the educational sector, researchers and educators are currently experimenting with elements of gamification to motivate and engage students. Freitas et al. developed a framework for identifying the gamification profile of students to tailor gamification projects to students' needs (2017, pp. 1, 8). They argue that the younger generation is used to modern digital technology and playing games, which is why the application of game elements in education seems to be a natural choice to increase engagement and motivation. Kuo et al. describe playing as the basic, intrinsic desire for relaxation (2017, p. 548). They proposed a gamified online platform to support classroom management. An extensive literature review analyzing existing studies about gamification and its application to software engineering education has been carried out by Souza, Veado et al. (2017, p. 170). The authors collected data of 106 studies from 1974 to 2016. They categorized the studies in terms of their gamification approach and knowledge area of software engineering. Their results showed that the most common approach, used in more than half of the studies, was game-based learning. Souza, Veado et al. describe game-based learning as "game applications that have defined learning outcomes" (2017, p. 171) and include serious games, edutainment and educational games. Approaches using game development-based learning were detected in 38 of the 106 studies, while gamification itself was found to be the least used approach, accounting for eight of all 106 studies (Souza, Veado et al., 2017, p. 178). The other studies used a hybrid approach. The most covered knowledge areas of software engineering were software process, software design, and professional practice. Software quality as well as software modeling and analysis were the least covered areas. This paper would most likely be categorized as a gamified approach addressing software modeling.

Souza, Constantino et al. introduced two elements of gamification, namely badges and leaderboards, into their software engineering course attended by 36 students at the Federal University of Minas Gerais in Brazil (2017, p. 276). They observed that students perceived badges very positively, while leaderboards caused a generally negative response. Azmi et al. reviewed existing literature in the field of gamification in programming education, arguing that traditional face-to-face techniques result in students losing motivation (2016, p. 112). They claim that computer science is a skill-developing subject and not a memorizing class, which is why programming classes should not be taught in a passive style, but rather actively engaging the students. Azmi et al. also interviewed experts from public universities in Malaysia and proposed to employ gamification to foster the usage of discussion forums for peer reviews (2017, pp. 1f). Wang et al. analyzed the effects of quizzes in the classroom, comparing paper quizzes, the non-gamified student-response system Clicker, and Kahoot!, a game-based "student-response system that has a main focus on student motivation and engagement through gamification" (2016, p. 730). They conducted an experiment at the Norwegian University of Science and Technology in an IT course with a total of 384 students. They were divided into three groups using different quiz tools. The results showed that students in the game-based group using Kahoot! reported a significantly higher engagement, motivation, concentration, and enjoyment (Wang et al., 2016, p. 736). This suggests that the gamified nature of Kahoot! provides additional value to quizzes and increases the active involvement of students. Furthermore, Wang and Lieberoth evaluated the impact of the audio and points features of a Kahoot! quiz in an experiment with a total of 593 students (2016, p. 738). The results showed that the audio and points features have a significant impact on engagement, motivation, concentration, and enjoyment. Chaiyo and Nokham compared the digital student-response systems Kahoot!, Quizizz, and Google Forms in an experiment at the Chiang Rai College in Thailand with 121 students (2017, p. 178). The results showed a significant difference in concentration, engagement, enjoyment, motivation, and satisfaction with Kahoot! receiving the best ratings in all categories (Chaiyo & Nokham, 2017, p. 181f). The students stated that a quiz "helped them to be aware of their level of knowledge and facilitates the understanding of the concepts and increases their learning process" (Chaiyo & Nokham, 2017, p. 182), suggesting that digital quizzes provide a good opportunity for students to self-evaluate their proficiency level.

Much earlier, Carl Rogers, the founding father of client-centered therapy and the person-centered approach (Rogers & Koch, 1959), considered self-evaluation an important part of learning. He stated that the "evaluation of one's own learning is one of the major means by which self-initiated learning becomes also responsible learning" (Rogers, 1983, p. 158). Therefore, he suggests that self-evaluation shall be included in approaches promoting experiential learning. However, person-centered learning is more than just the application of person-centered methods of learning: it is about the attitude of the teacher which allows him or her to become a facilitator. Rogers describes this attitude to promote learning and enhance the climate for learning as a transparent realness including pricing, caring, trust, and respect for the learner (Rogers, 1983, p. 133). Studies confirmed that students learn more when they receive a high level of understanding, caring, and genuineness (Rogers & Freiberg, 1994, p. 253). Motschnig et al. introduced learner-centered principles into an undergraduate course on human-computer interaction attended by 200 students, which was rated the best computer science course at bachelor's level assessed by a cohort of more than four students (2016, p. 1). The course concept included multiple occasions for interactions such as asking questions, giving feedback, and two open book multiple-choice tests. Standl compiled a catalog of person-centered design patterns for computer science education at secondary school level, including the self-check as one method of person-centered learning (2013, p. 120). However, Standl's self-check focusses on written feedback on projects, raising the question whether the pattern can be generalized for closed questions and smaller exercises.

Due to the closed nature of multiple-choice questions, they are said to only test lower levels of thinking (Scully, 2017, p. 4), usually referring to Bloom's Taxonomy of Educational Objectives (Bloom, 1956). Bloom's original taxonomy consisted of six classes: knowledge, comprehension, application, analysis, synthesis, and evaluation (Bloom, 1956, p. 18). As the right answer is always provided as a response option, critics argue that multiple-choice questions are incapable of assessing cognitive processes beyond recall and recognition (Scully, 2017, p. 4). However, Scully argues that the potential of multiple-choice questions assessing higher levels of thinking may be underestimated due to the overrepresentation of multiple-choice questions only assessing lower levels of thinking. She conducted an extensive literature review in this field and gave recommendations on how to write multiple-choice questions for higher levels of thinking. Scully argues that multiple-choice questions, if designed accordingly, are also able to assess at the levels of application and analysis (Scully, 2017, p. 10).

This paper analyzes the combination of the introduced concepts in the higher education sector. More specifically, it evaluates the applicability of gamified single-choice questions using Kahoot! as self-evaluation tool in a flipped, person-centered course on the introduction to programming. The research questions are: (1) What additional value do game-based student response systems provide to a flipped classroom setting in higher education? and (2) Do single-choice questions have the potential to promote higher-level thinking? Moreover, based on our experience, we derive good-practice guidelines for integrating game-based student response systems into flipped classrooms.

Case Study Framework

A comparative case study design was chosen as the overall framework for this research. For this matter, an introductory bachelor-level course on C programming attended by 28 students was modified in the winter term of 2017/18 and qualitatively and quantitatively compared to the previous year. The following data of the case study was analyzed to evaluate the effects of the introduced changes with the permission of the students:

Course feedback: after the final exam, both in 2016/17 and 2017/18 the students filled out an anonymous final feedback form consisting of standardized questions. Before they gave written feedback on what they liked about the course and what could be improved about it, they were asked to rate their agreement with statements about the course, the instructor, and the students (A1-A9) on an ordinal scale from "agree"* [1] to "disagree"*.

Online questionnaires: after a quiz was held in class in 2017/18, the students voluntarily filled out an anonymous online questionnaire consisting of several statements (B1-B7) to be rated on an ordinary scale. Moreover, students were able to give written feedback on what they liked and what they suggested for improvement.

Quiz results and grades: the scores on the twelve Kahoot! quizzes and the number of correct answers to the 171 single-choice questions as well as the students' grades on the exam were analyzed.

The following table shows the modified and unmodified variables of the course as well as the linkage to the evaluation instruments aiming to assess the respective course aspect:

Course aspect	**Original course in 2016/17**	**Introduced changes in 2017/18**	**Evaluation instrument**
Course structure and content	Flipped classroom setting in a part-time degree program; weekly face-to-face lessons of 90 minutes; predefined study schedule; introduction to C programming	No changes introduced	A1: "I was properly informed about goals, content, methods, and performance requirements of the course"* A2: "The course was well structured regarding content"* A7: "The course design was interesting and inspiring"* A9: "Overall, I'm very satisfied with the course"*
Course material	Full-text scripts covering course material	Kahoot quizzes added as course material	A3: "The used material (e.g. books, scripts, worksheets) reasonably supported the learning process"*
Recapitulation	Weekly graded 10-minute written assignments; oral in-class recapitulation; obligatory	Weekly Kahoot! quizzes; one extra credit for the three best participants; voluntary	B1: "I have learned something new in today's class"* B2: "The subject matter was well consolidated through the repetition"* B3: "Due to the quiz, I was more engaged in the learning process than usual"* B4: "The quiz motivated me to self-study more than an oral in-class repetition"*

[1] Items marked with * were translated from German.

			B5: “Overall, I am satisfied with today’s quiz”* B6: “I would like to recapitulate the next chapter using a quiz as well”* B7: “I would like to have this number of teaching units utilizing a quiz:*”
Exercises	Weekly submissions; 0-24 points received per submission; sum of points used for grading; feedback within one week	Presentations awarded the grades of 1 (“+”), 0 (“~”), or -1 (“-“); average for presentations used for grading; immediate feedback	A4: “There were diverse possibilities to actively process the content, e.g. exercises, discussions, projects, or small-group work”* A8: “The students actively contributed to the learning process as an audience and in smaller groups, e.g. with own examples and questions”*
Project	Final complex project requiring students to implement a simple hash table	No final project	A5: “The practical relevance was made clear”*
Exam	One final exam	Midterm exam covering two learning outcomes and final exam covering three	Exam results
Grading	A minimum of a 50% score required on the weekly 10-minute assignments, weekly submissions, final project, and final exam each; weighing: 25% assignments, 25% submissions, 18.75% project, 31.25% exam	A minimum of a 50% score required on each of the five learning outcomes and on the average for all presentations; weighting: 50% exam, 50% presentations; additional 0.5% per extra credit on the quizzes	Quiz results and grades
Instructor	4 years’ experience in teaching; person-centered attitude	Same instructor; one more year of experience in teaching	A6: “The course instructor contributed to a good learning atmosphere”*

Table 1: Comparative case study setup

To summarize, the grading aspects were reduced to two items and the way of recapitulation was modified by introducing Kahoot! quizzes, while the overall course structure, content, and schedule stayed the same. A strict study schedule, regular teacher grading, and exams are rather characteristic for traditional classrooms, but were required due to administrative constraints and the fixed curriculum. However, the teacher applied person-centered principles to the course in various ways. First, the students were included in the decision on how to recapitulate the content studied at home. The same teacher conducted a Kahoot! quiz with the students in the preceding course, which was highly appreciated by the students. A discussion showed that the students wanted to have regular in-class Kahoot! quizzes in the next course, which is why the teacher prepared quizzes recapitulating the respective content studied at home. Second, the exams were designed in a practical way, aiming to assess higher levels of thinking. Third, the teacher approached the students with a person-centered attitude and reacted to their feedback. For instance, when the students expressed their wish to take a mock exam a week before the midterm exam, the teacher prepared another exam as a sample and simulated an exam situation.

In total, twelve quizzes consisting of 171 single-choice questions were prepared and conducted. The Kahoot! quizzes were not used for grading to make them more enjoyable. They were designed to assess different levels of thinking: they used a mixture of questions aiming to evaluate knowledge, comprehension, application, and analysis. An example for a question at the *knowledge*-level would be “Which of the following is not a standard datatype in C?”*, providing the four answers string, double, int, and float.

At the level of *comprehension*, asking students to choose the correct variable declaration could verify whether students understand the basic concept of declaring new variables. For instance, providing a correct declaration such as “double fa[256];” as well as an assignment operation, a function declaration, and a function call could verify whether students comprehend the rules of variable declaration and its distinction from related concepts. Flipping the items, as suggested by Scully (2017, p. 6), is also an option.

The *application* of a concept requires real-life scenarios as students are asked to demonstrate that they can apply an abstraction in appropriate situations (Bloom, 1956, p. 120). For instance, the students were given a stimulus showing a small program including a variable definition with initialization, a bit operation, several conditional statements, and output statements. While it is enough at the comprehension level to understand the principles of those concepts, application requires students to make use of knowledge in a given situation. They need to be able to distinguish the presented statements, apply their specific rules to the particular situation one after another.

The *analysis* level requires a breakdown of material into parts and detection of their relationship and organization (Bloom, 1956, p. 144). Questions designed for the analysis level included a stimulus of a rather complex practical example. For instance, the students would be asked to analyze a given program consisting of multiple functions, decide whether the program compiles, and, if so, identify the correct output. First, such a question requires the students to break down the program into parts, such as include statements, function prototypes, implementations, local and global variables, pass-by-value parameters, pass-by-pointer parameters, assignments, loops, conditional statements, and operations. Second, it presents the need to identify the relationships between those concepts and decide whether the presented structures match each other, e.g. whether a function call matches its signature, as well as the rules students have learned. Questions on the analysis level could also ask students to identify the error in a rather complex program or decide which of multiple shown programs solve a given problem.

Findings

Course Feedback

In 2017, the voluntary and anonymous feedback form was filled out by 19 of 19 present students after the exam, while in 2018, 19 out of 20 students filled it out in 2018. The collected data was mapped to a scale from 1 ("disagree"*) to 4 ("agree"*) and averaged for the sake of simpler visual comparison.

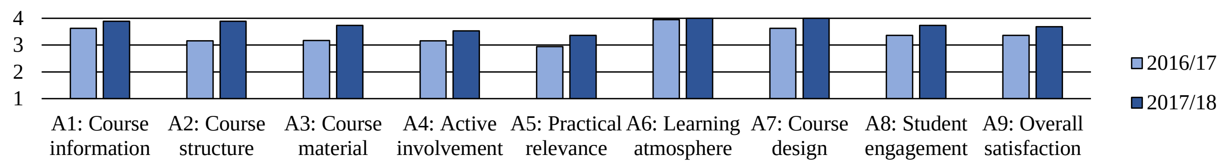

Figure 1: Final course feedback

(Fig. 1) shows the results of the final course feedback in the winter term of 2016/17 and 2017/18. An overall increase in agreement to every single of the positive statements can be observed. Interestingly, the greatest increase was observed regarding the course structure with an average improvement of 0.74. This difference proved significant in a nonparametric, continuity corrected, one-sided Wilcoxon rank sum test with the alternative hypothesis that the agreement to this statement was greater in 2017/18, in which the new course modalities were introduced. With a p-value of $3.8*10^{-4}$, we accepted the alternative hypothesis. Considering that the basic course structure, the underlying study concept, and the content have not changed, the fact that the students found the course better structured in 2017/18, when quizzes were used, is rather surprising. However, this matches the findings of Chaiyo and Nokham, who discovered that quizzes help students to be aware of their knowledge level and to better understand the concepts (2017, p. 182). It seems that the regular quizzes provided students a good framework and an overview of the content.

The second highest difference was measured in the perceived suitability of the course material to support the learning process: the average level of agreement with the respective statement could be increased by 0.57 with a p-value of $7.3*10^{-3}$, assessed via a one-sided Wilcoxon rank sum test with the alternative hypothesis that the agreement to this statement has increased. As the quizzes were the only changed aspect regarding course material, students therefore seem to appreciate the regular quizzes and see them as viable course material supporting self-study. Upon request of the students, the teacher provided them with the quiz questions as Kahoot! challenges before the exams, enabling them to play the quizzes again at home.

The course design was rated notably better in 2017/18 with an increase of 0.37 and a p-value of $4.6*10^{-3}$ in a one-sided Wilcoxon rank sum test with the alternative hypothesis that the agreement to the positive statement about the course design has increased. Considering that the basic course design was not changed, the students therefore appeared to perceive a course utilizing quizzes as more interesting and inspiring than regular courses. A written response to the question about what students liked about the course was: "Quiz! Motivates to actively cooperate, diffuses the tension, many things get clarified/explained. TOP! 1000 times better than boring oral in-class recapitulation"*. Overall, 12 out of the 19 students mentioned the quiz as something they liked about the class. This illustrates the students' appreciation of the quiz experiment. A one-sided Wilcoxon rank sum test with the alternative hypothesis that the agreement to the statement was greater in 2017/18 showed a slight improvement in overall satisfaction of 0.32 with a p-value of 0.047.

Although the average level of agreement to all other positive statements about the course increased, the respective alternative hypotheses could not be proven significant with the given sample size. Most notably, in the winter term 2017/18, all of the 19 students fully agreed with the statement "The course instructor contributed to a good learning atmosphere"*. Despite the difficulties of learning to program, including negative grades on the exam, the students valued the atmosphere in class. The teacher used a person-centered approach and tried to establish a climate of trust, respect, and understanding. Comments on what students liked about the course included "the instructor", "relaxed atmosphere"*, "good style, patient"*, and "friendliness"*.

A correlation analysis using Spearman's correlation coefficient was carried out to identify the relationship between the variables. In both years, the highest correlation was observed between the overall satisfaction and suitability of the course material with $r_s=0.76$ in 2017/18 and $r_s=0.62$ in 2016/17 respectively. The test for correlation showed a linear dependence with a p-value of $1.8*10^{-4}$ and $6.1*10^{-3}$ respectively. This supports the hypothesis that quizzes provide additional value to the course as the perceived suitability of the course material showed a significant increase. In 2017/18, the second highest correlation was measured between practical relevance and active involvement with $r_s=0.67$ ($p=1.8*10^{-3}$), which was not observable the year before ($r_s=0.09$, $p=0.72$). One possible explanation could be that the new course modalities made the connection between the content and practical relevance clearer. In 2016/17, the second highest correlation was found between the course structure and practical relevance with $r_s=0.55$ ($p=0.014$). This was not the case in 2017/18: the correlation was negative with $r_s=-0.35$ ($p=0.15$).

Online Questionnaires and Quiz Results

The following diagrams show the results of the voluntary questionnaires collected right after a quiz was conducted. The number of responses on the questionnaires after the twelve quizzes (Q1-Q12) were: 27, 19, 14, 15, 17, 8, 14, 5, 10, 12, and 10. Due to time constraints, two quizzes were conducted in the last unit (Q11 and Q12).

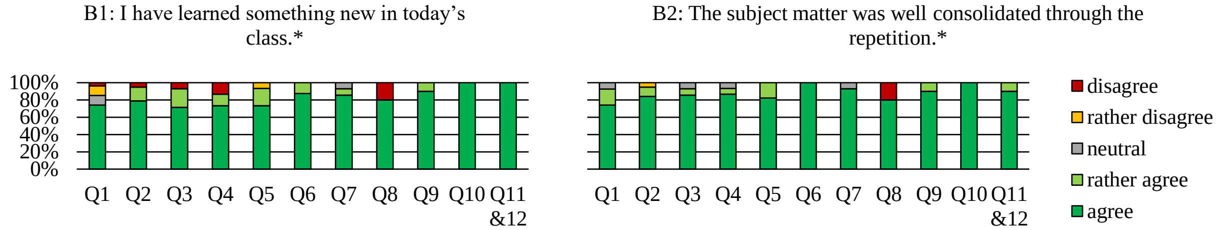

Figure 2: Learning effect

(Fig. 2, left) shows the students' agreement with the statement B1 on a scale from "disagree"* to "agree"* in all eleven lessons in which a quiz was conducted. The reported learning effect remained steady and did not show a significant drop over the semester: at least 74% of all students (rather) agreed with this statement after each quiz. In about half of the quizzes, one outlier could be observed who tended to disagree with all statements. A deeper analysis reveals a rather contradictory behavior: in each iteration, the outlying record stated that this class participant wants to have a quiz in 100% or 75% of all the units. This could be an indication that this person inverted the scale and wanted to give the best ratings rather than the worst. Another student explained the disagreement differently: "[…] because I have known the topic of this unit for years"*. Written answers to what they liked about learning effect were: "that after each point it is explained what the program does"*, "tricky questions"*, "the tricky questions with pitfalls"*, "that there are also tricky questions, you learn more though that"*, "the questions and the explanations"*, "explanation of the answer after each question"*, "the solution to the question + explanation"*, "good explanations for better understanding"*, and "questions that are hard to understand get explained well"*.

In addition, the students were asked to rate how well the subject matter was consolidated through the quiz. A similar distribution of the answers could be observed, as depicted in (Fig. 2, right). In each iteration, 80% of the students or more (rather) agreed with this statement. The outlying records were rather neutral here, except for the eighth quiz. Essentially, the positive trend remained steady until the end of the course.

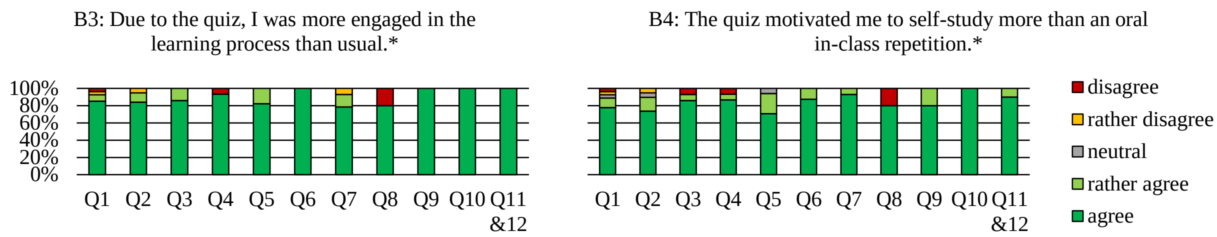

Figure 3: Student engagement and motivation

(Fig. 3, left) shows the self-reported student engagement in each iteration. If the outlier is not taken into consideration, only one student rather disagreed with this statement once, which happened in the first iteration. After that, every student except for the outlier rather agreed that they were more engaged in the learning process than usual. Written responses concerning student engagement on the question what they liked include "active contribution and the discussion of wrong answers"* and "the possibility to contribute"*. These results confirm the findings of the literature that students can be more engaged and actively involved using elements of gamification in class.

(Fig. 3, right) compares the students' motivation when using quizzes with orally recapitulating the content in class. The agreement to this statement remained steady and did not experience a notable drop over time. If the outlier is not considered, only one respondent rather disagreed with this statement. Written answers about what the students liked about the quizzes highlighted the pleasure of doing them, including: "the stimulation"*, "it is fun"*,"the stimulation and motivation you get from the quiz"*, "that it is fun"*, "fun during recapitulation"*, "the race motivates to study"*, "competition feeling, fun"*, "that you can compete with your peers"*, "you have fun and study at the same time"*, and "fun to study"*. These ratings and written responses show the students' appreciation of the quiz experiment and suggest that gamified student-response systems have indeed a positive impact on the motivation of students to study.

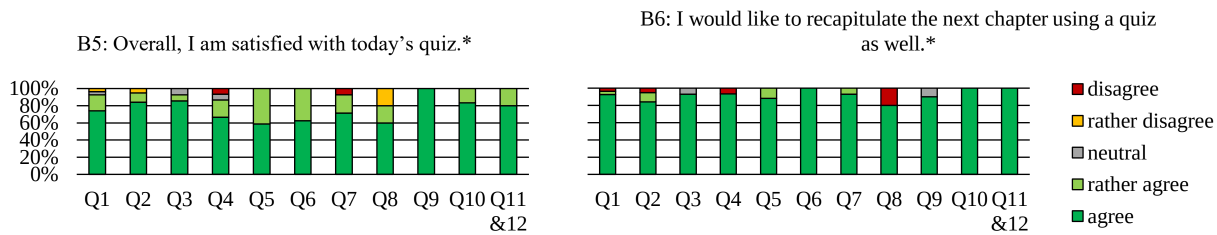

Figure 4: Satisfaction and weekly voting

The students' satisfaction with the quizzes is shown in (Fig. 4, left). The students seemed to be satisfied with the quizzes as no respondent (rather) disagreed with this statement in each iteration except for a recurring outlier. The continuous positive feedback was the reason that the instructor took up the quizzes as a regular element of the course schedule.

The students were included in the decision-making process whether the quiz experiment should be continued by voting whether the next unit should also employ a quiz. The results of this weekly voting are depicted in (Fig. 4, right). The students were in almost full agreement that the quiz experiment should be continued. Except for the outlier with a possible inverted scale, not a single student disagreed with this statement.

The students were also asked for an estimate of how many of the face-to-face classes should employ a quiz every week: 100%, 75%, 50%, 25%, or 0%. A quiz was held in eleven of the 14 units, which is why "75%" reflected the situation best. Only six of the 151 responses wished for less than 75% lessons employing a quiz. The average over all iterations is 94.5%, which again shows the students' appreciation of quizzes in class. The students were asked for a rating in each class to observe whether the preferred number drops over time. However, the opposite took place: while the overall average was 87% in the first lesson, students wished for a mean of 97.5% lessons employing a quiz at the end of the semester. This again underlines the suitability of quizzes from a student's point of view.

The percentage of correct answers to all twelve quizzes was 72%, 67%, 60%, 68%, 65%, 64%, 50%, 74%, 48%, 70%, 68%, and 65%. The quizzes were therefore challenging and required students to understand given stimuli, apply their knowledge and analyze given programs. The midterm and final exams were of a practical nature, asked students to write short programs, understand given pieces of code, and identify correct outputs, aiming to assess higher levels of thinking. If the sum of the two exam grades is compared to the exam grades of the previous year, an overall increase by a factor of 12% can be observed. This suggests that carefully prepared quizzes promote higher-level thinking. However, this relationship could not be proven significant with the given sample size.

Discussion and Interpretation

We have found the following answers to our research questions based on the evidence and our observations:

(1) What additional value do game-based student response systems provide to a flipped classroom setting in higher education? In comparison to the preceding class that did not use Kahoot!, the students of 2017/18 perceived the course structure and design to be considerably better, although the basic structure of the course itself did not change. This suggests that regular quizzes help students understand the structure of the course content, possibly because of the regular self-evaluation. Furthermore, the course material used to support the learning process was rated significantly better than the year before, which indicates that students appreciate the quizzes as valuable course material indeed. The suitability of course material for the learning process was found to correlate with the overall satisfaction with the course in both years, which may be the reason the satisfaction also significantly increased. The regular online questionnaires showed that almost all students reported that they liked the quizzes, experienced a higher learning effect, stayed more engaged, were more motivated, and would like to have a quiz in about 87% of all face-to-face classes. These effects did not experience a notable drop during the progression of the course.

(2) Do single-choice questions have the potential to promote higher-level thinking? The grades on the exam aiming to assess higher-level knowledge were improved by 12%. However, this effect could not be proven significant with the given sample size. Further research is necessary to answer this question. Our observations showed that students analyzed programs and applied their knowledge to given situations, which points to cognitive processes higher than the levels of knowledge and comprehension.

We are aware of certain limitations of this case study. First, single-choice questions always include the correct answer, which is why it can never be guaranteed that the cognitive processes happen at a higher level for each student. Random guessing certainly does not promote higher-level thinking. Second, multiple variables were changed at the same time as the course was optimized in multiple ways. However, the changes only addressed aspects of assessment and grading, while the basic structure and content of the course remained the same. Third, 28 students are a rather small sample size, which is a common problem in educational research. Finally, the drop in the number of responses on the online questionnaire may imply an introduced bias. As the feedback was voluntary, such a drop is reasonable and common. However, the drop is partly explainable by the dropout of students: the average response rate after a quiz was 70%. Despite these limitations, we are confident that our findings are of value for researchers and educators in the field of computer science education.

Recommendations

Based on the collected data, our observations, and experience, we propose the following recommendations for the use of game-based student response systems in flipped classrooms in the field of computer science education:

Use diverse question types: the questions of quizzes should aim to assess different levels of thinking. Knowledge- and comprehension-based questions can usually be quickly and easily answered, enabling students to experience a sense of achievement and “speeding up” the quiz. However, they can quickly lose their appeal and do not promote higher-level thinking. Questions aiming to assess the application and analysis level of thinking need more time to be answered, slow down the quiz and exhaust students. We therefore suggest alternating between these two basic question types, starting with questions on the lower levels to build up tension, and complementing the learning effect with a few higher-level questions afterwards. When students get tired, it is time for the next set of quick lower-level questions to bring back tension. We also found that negated questions require additional cognitive effort and cause students to make mistakes. We therefore recommend using them sparingly or avoiding them completely.

Focus on explanations: the main goal of in-class quizzes is not to entertain students, but rather to enrich the learning process. Therefore, it is important to explain why an answer is correct or incorrect, especially if there is a high number of incorrect responses. This provides a good opportunity for the teacher to identify difficult topics that need further explanation. Higher-level questions may require special attention and can be discussed by going through the respective program step-by-step. Asking the students to explain why an answer is correct is a good way to include them in the learning process and provides them with additional explanations.

Choose high-quality distractors: the quality of the distractors is essential to enable higher-level thinking. If the correct answer is obvious, students do not need to analyze a program to distinguish the responses. Including answers based on common misconceptions and edge cases can help discuss important aspects of programming.

Short quiz duration: interactive quizzes activate students and require them to pay special attention. As this constant tension is demanding, students get exhausted quickly, which is why the duration of a quiz should be chosen accordingly. Our experience showed that quizzes should take no longer than 30 minutes as students start to get tired.

Answer questions first: we found that answering students' questions and discussing exercises first was a good approach as it gives every student a fairer chance to succeed in the quiz. In this way, students who had problems with understanding a topic studied at home are granted an opportunity to have their questions answered. Even if some quiz questions get answered during the discussion, this does not lessen the quizzes' learning effect because not every student chooses the right answer to the same question. Not only does this provide the teacher a possibility to self-evaluate his or her explanation, but also rewards students who paid attention.

Decide whether to use audio: the students were included in the decision-making process whether audio would be used in each subsequent class. Sometimes they voted to use audio, sometimes they voted against it. We found that audio provided additional excitement and tension, which may not always be beneficial: if students are having an exhausting day, they might prefer a quieter session. Therefore, we recommend including students in this decision.

Provide questions as course material: based on our experience, students grow fond of the quizzes, which is why they will sooner or later ask for the quiz questions to help them study for the exam. We recommend supporting them by exporting the questions and providing them as a document after every quiz or using Kahoot! challenges.

Decide grading strategy: we did not include the Kahoot! quizzes in the evaluation scheme of the class. Although some educators use Kahoot! quizzes to grade their students, we decided against this for various reasons. First, we believe that grading takes most of the fun out of student response systems. Second, it makes the course dependent on the infrastructure and availability of the website. Some students disconnected due to a lost internet connection or a crashed device. In order to be fair, one would have to build a rather complex backup strategy to avoid frustration among students in case of a disconnect. Third, if the quizzes were used for grading, privacy and security concerns arise as Kahoot! stores certain data.

Conclusion

The feedback on the quiz experiment was very positive, which is why we consider it a success and see gamified quizzes as a valuable addition to engineering courses. The final course feedback was improved in every aspect, while we constantly received overwhelmingly positive feedback each week a quiz was conducted. The students appreciated quizzes as part of the course and saw it as valuable study material. The quizzes helped them understand the course structure and they found the course more interesting and satisfying. Further work includes the suitability of single choice questions to support higher-level thinking, which could not be proven with this sample size.

At the end, we wish to note that we think that an essential success factor was the fact that the instructor listened to students and endeavored to see the course through their eyes. The quizzes were just one apparent (may we say "cool") instance to materialize this listening – they were never imposed and always asked for by students whose voice was heard. They had an active part in designing their course, which signals caring and respect, one essential characteristic of person-centered learning. Moreover, the answering of questions in class and the transparent and immediate feedback that the quizzes provided signal openness and transparency – another core ingredient of a person-centered classroom. We highlight listening, caring, and transparency due to the importance of communicating these attitudes to students in person-centered learning. We tried to react to the wishes of our students during the quiz experiment. For instance, students wished for a special tagging of "tricky" questions to signalize a pitfall, which we incorporated in further quizzes. We believe that this is our most important recommendation: listen to your students as individuals to facilitate learning.

Acknowledgements: We thank the students for participating in the quiz experiment and their willingness to share their feedback and experiences, enabling us to conduct our research in the field of engineering education.

References

Azmi, S., Ahmad, N., Iahad, N. A., & Yusof, A. F. (2017). Promoting students' engagement in learning programming through gamification in peer-review discussion forum. In *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)* (pp. 1–6). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICRIIS.2017.8002543

Azmi, S., Iahad, N. A., & Ahmad, N. (2016). Attracting Students' Engagement in Programming Courses with Gamification. In *2016 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)* (pp. 112–115). Piscataway, NJ: IEEE. https://doi.org/10.1109/IC3e.2016.8009050

Banville, L. (2016). Education Gamification Markets Expected to Grow Globally. Retrieved March 13, 2017 from http://www.gamesandlearning.org/2016/12/14/gamification-elements-continue-to-draw-dollars/

Bergmann, J., & Sams, A. (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day*. Arlington, VA: International Society for Technology in Education.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. *Taxonomy of Educational Objectives: The Classification of Educational Goals*: D. McKay.

Chaiyo, Y., & Nokham, R. (2017). The effect of Kahoot, Quizizz and Google Forms on the student's perception in the classrooms response system. In *International Conference on Digital Arts, Media and Technology (ICDAMT)* (pp. 178–182). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICDAMT.2017.7904957

Freitas, S. A. A., Lacerda, A. R. T., Calado, P. M. R. O., Lima, T. S., & Dias Canedo, E. (2017). Gamification in education: A methodology to identify student's profile. In *2017 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). Piscataway, NJ: IEEE. https://doi.org/10.1109/FIE.2017.8190499

Kanelopoulos, J., Papanikolaou, K. A., & Zalimidis, P. (2017). Flipping The Classroom to Increase Students' Engagement and Interaction in a Mechanical Engineering Course on Machine Design. *International Journal of Engineering Pedagogy*, *7*(4), 19–34. https://doi.org/10.3991/ijep.v7i4.7427

Kumar, J. M., & Herger, M. (2017). *Gamification at Work: Designing Engaging Business Software*. Aarhus, Denmark: The Interaction Design Foundation.

Kuo, M.-S., Chuang, T.-Y., Tao, S.-Y., & Yang, J.-W. (2017). Designing a Digital Gamification Platform to Support Classroom Management. In *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*. Piscataway, NJ: IEEE. https://doi.org/10.1109/IIAI-AAI.2017.69

MarketsandMarkets Research. (2016). Gamification Market by Solution & Applications Type. Retrieved March 13, 2017 from https://www.marketsandmarkets.com/Market-Reports/gamification-market-991.html

Motschnig, R., Sedlmair, M., Schroder, S., & Moller, T. (2016). A team-approach to putting learner-centered principles to practice in a large course on Human-Computer Interaction. In I. F. i. E. Conference (Ed.), *The crossroads of engineering and business: Frontiers in Education 2016 : October 12-15, 2016, Bayfront Convention Center, Erie, PA* (pp. 1–9). Piscataway, NJ, USA: IEEE. https://doi.org/10.1109/FIE.2016.7757576

P&S Market Research. (2016). Gamification Market. Retrieved March 13, 2017 from https://www.psmarketresearch.com/market-analysis/gamification-market

Research and Markets. (2016). Global Education Gamification Market 2016-2020. Retrieved March 13, 2017 from https://www.researchandmarkets.com/reports/3985296/global-education-gamification-market-2016-2020

Rogers, C. R. (1983). *Freedom to Learn for the 80's* (2nd). Columbus, OH: Merrill Publishing.

Rogers, C. R., & Freiberg, J. H. (1994). *Freedom to Learn* (3rd). Columbus, OH: Merrill Publishing.

Rogers, C. R., & Koch, S. (1959). A Theory of Therapy, Personality and Interpersonal Relationships as Developed in the Client-centered Framework. *Psychology: a Study of a Science*, *3*, 184–256.

Scully, D. (2017). Constructing Multiple-Choice Items to Measure Higher-Order Thinking. *Practical Assessment, Research & Evaluation*, *22*, 1–13.

Souza, M. R. d. A., Constantino, K., Veado, L., & Figueiredo, E. (2017). Gamification in Software Engineering Education: An Empirical Study. In *2017 IEEE 30th Conference on Software Engineering Education and Training (CSEE&T)* (pp. 276–284). Piscataway, NJ: IEEE. https://doi.org/10.1109/CSEET.2017.51

Souza, M. R. d. A., Veado, L., Moreira, R., Figueiredo, E., & Augustus Xavier Costa, H. (2017). Games for Learning: Bridging Game-Related Education Methods to Software Engineering Knowledge Areas. In *2017 IEEE/ACM 39th*

International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET) (pp. 170–179). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICSE-SEET.2017.17

Standl, B. (2013). Conceptual Modeling and Innovative Implementation of Person-centered Computer Science Education at Secondary School Level (Doctoral thesis). University of Vienna, Vienna, Austria.

Wang, A. I., & Lieberoth, A. (2016). The effect of points and audio on concentration, engagement, enjoyment, learning, motivation, and classroom dynamics using Kahoot! In *10th European Conference on Game Based Learning (ECGBL)* (pp. 738–746). Sonning Common, England.

Wang, A. I., Zhu, M., & Sætre, R. (2016). The Effect of Digitizing and Gamifying Quizzing in Classrooms. In *10th European Conference on Game Based Learning (ECGBL)* (pp. 729–737). Sonning Common, England.

Zichermann, G., & Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. *O'Reilly Series*. Sebastopol, CA: O'Reilly Media.