Improving Students' Motivation and Focus Through the Gamification in the Computer Science Peer Instruction Methodology (CSPI)

Ricardo Caceffo and Rodolfo Azevedo

Institute of Computing, University of Campinas (Unicamp)
{caceffo; rodolfo}@ic.unicamp.br

Abstract: The Peer Instruction (PI) is an active learning, student-centered approach designed to engage students through the administration of conceptual questions. On its turn, the CSPI is an adaptation of the PI to the computer science area, mainly Introductory Programming (CS1) courses. In this work, we present the initial assessment of the gamification adoption into the CSPI. Two CSPI/CS1 courses were administered, one without gamification (control experiment) and one with it. We identified that, although participants originally didn't expect that gamification would affect their experience, results suggest that gamification positively impacts participants' motivation and focus. Future work involves the analysis of collaboration and interaction variables through learning analytics tools.

Keywords: *Gamification; active learning; peer instruction; Computer Science peer instruction; CS1; assessment; motivation.*

1. Introduction

In contrast to the traditional instructor-focused classes, in 1991, Bonwell & Eison [1] explained that students must do more than just listen: they should read, write, discuss, and be actively involved in solving problems, which is known as Active Learning.

One of the most adopted Active Learning approaches, the Peer Instruction (PI), was born in the physics area. As described by Kelly, M. [3], the PI is a student-centered teaching method designed to engage students through the administration of conceptual questions, which help resolve misconceptions related to the class content.

In a typical PI class, the theoretical content is interspersed with questionnaires, usually multiple-choice. As explained in [3], students have two opportunities to answer each question, the first after a round of individual reflection and then again after a discussion round with a peer. Based on the number of right answers, the instructor can make adjustments to address students' issues.

We have designed a PI adaptation to the scope of computer science, mainly focusing on CS1 courses [4], and defined as Computer Science Peer Instruction (CPSI). In the same way as the PI, at some point in a CSPI class, the instructor administrates a set of multiple-choice questions to the students. The main difference is that, on the CSPI, students have only one chance to answer each question. Initially, students work alone for 1 minute, having access only to the question statement. Then, they work in groups of 2 or 3 students, having 2 minutes to think together and select one of the available choices for the question. The instructor can extend the time if required. Finally, the instructor, considering the number of right answers, explains the solution.

In both PI and CSPI, technology can support the automatic submission, correction, and display of the responses. In the current study, we adopted the clickers (iClicker brand [6]), small remote-control devices that have 5 buttons (letters from A to E). The clicker device can submit in real-time the answer (letter) pressed by a student, which is received by a base connected to the instructor's notebook. A software then shows a consolidated chart of answers from all participants.

On its turn, as described by Klock *et al.* [8], the concept of gamification relates to a set of techniques that extend game design elements to other contexts, like electronic commerce and learning environments. Related to the latter, the authors state that the adoption of gamification *"is justified as a way to motivate and engage students"* [8].

Moreover, Feldbusch *et al.* [2] explain that gamification is a wide approach, supported by the adoption of a variety of game elements, like a scoring system, badges, achievements, level systems, and so on.

On this work we are focusing on the competition aspect, in a PI learning environment, supported by the adoption of a public score system. The contributions are the initial assessment of the inclusion of the gamification approach into the CSPI. We want to identify whether gamification is a positive factor and whether it can affect the students' motivation, focus, and social collaboration.

2. Methodology

Lazar *et al.* [6] explain that *experimental research* normally starts from a preformed theory, usually organized as one or more hypotheses, which supports the design of experiments to collect and analyze data that will ultimately prove (or not) the original hypothesis. The cycle then repeats itself, being the hypothesis updated from the previous results, leading to new experiments, data collection, and so on. Figure 1 shows a diagram, based on Lazar experimental methodology [6, pp-305, Figure 11.1], illustrating how this research is organized.

On this research, preformed theory (Figure 1, item 1) was organized as the following **Research Questions:**

- **RQ1:** Does gamification affects students' motivation and focus in CSPI classes?
- **RQ2:** Does gamification affects students' social interaction and collaboration in CSPI classes?
- **RQ3:** Overall, is gamification a positive factor to support CSPI classes?
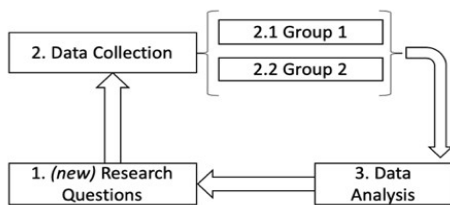
Figure 1. Experimental research methodology adopted in this study. Based on Lazar [6, pp-305, Figure 11.1].

Then, the **Data Collection** step (Figure 1, item 2) was organized as two experiments, defined as Groups G1 and G2. The Group G1 experiment was held at the Programming Week (PW), an annual event organized by the Brazilian Computer Society (BCS). In the PW, the *"Introduction to the Computational Thinking in Python"* course was taught by one of the researchers using the CSPI methodology. Participants (N=6) were all K-12 instructors.

The course had a total of 25 hours, with 20 "theoretical" hours in the classroom, in which theory content (*e.g.* PowerPoint slides) were interspersed with multiple-choice questions that should be answered through the clickers system. Participants were divided into 3 pairs, sharing one clicker per pair. The other 5 hours were freely used by the participants to solve problems in the laboratory. These lab hours were supervised by a TA. The CSPI methodology was not employed in the laboratory classes.

In the course's last day, participants from the Group G1 took a survey composed of 4 Likert scale questions (questions Q1 to Q4, as presented in Table 1) and an open-ended question, in which they could freely express themselves about the course, CSPI methodology, usage of the clickers, *etc.*

Table 1. Question statements (Likert scale questions).

Q1	*I would be more motivated to participate and use the clickers if it were adopted the idea of gamification.*
Q2	*I prefer competitions with prizes.*
Q3	*I usually talked to my colleagues when solving the exercises.*
Q4	*I liked the classes with clickers and I would like the teacher to continue with this approach in the future.*
Q5	*Gamification was a positive approach to increase students' motivation and focus.*

The Likert scale questions were designed to map the research questions in the following way: RQ1 (questions Q1 and Q2); RQ2 (question Q3) and; RQ3 (question Q4).

On its turn, the Group G2 experiment was held at a PW event at a different year than the Group G1. Participants (N=8) were all K-12 instructors, except by one, that was a K-12 student. The course organization was the same as the Group G1, except by the number of participants pairs (4) and by the adoption of a modified CSPI gamified version.

The CSPI gamified version was the same as the traditional CSPI methodology, except by the fact that the instructor added a score table in the blackboard. The table was used to publicly record the scores for each group, being updated after students answered the questions through the clickers. As the submission is anonymous, the instructor publicly asked which groups had correctly answered the question.

In the same way as Group G1, in the course's last day participants from Group G2 took a survey composed of 4 Likert scale questions (questions Q2 to Q5, as presented in Table 1) and an open-ended question. The Likert scale questions mapped the Research Questions in the following way: RQ1 (questions Q2 and Q5); RQ2 (question Q3) and RQ3 (question Q4).

The **Data Analysis** step (Figure 1, item 3) was conducted considering the *quantitative* data extracted from the 5-point Likert scale questions' answers and the *qualitative* data extracted from the open-ended questions' answers. Results from the Likert scale questions were initially converted to the numbers 1 (strongly disagree) to 5 (strongly agree), respectively. Then, for each question, it was calculated the *mean* of these values.

The next step was to assess whether the differences from both groups (G1 and G2) were statistically significant. We decided to adopt the nonparametric Mann-Whitney U test, as the data was collected from two independent groups of participants [6]. As explained by Nachar [7], the Mann-Whitney U test null hypothesis (H_O) stipulates that the two independent groups are homogeneous and have the same distribution. The nonparametric tests results are presented in subsection 3.3.

Finally, closing the cycle described by Lazar [6], the research follow-on (Figure 1, item 1) is presented in the discussion section (Section 5).

3. Results
3.1 G1 Group (without gamification)
Table 2 presents the results from the Likert scale questions related to the G1 Group:

Table 2. N= 6 participants. *SD* stands for "Strongly Disagree"; *D* for "Disagree" and so on. Questions are presented in Table 1.

	SD	*D*	*N*	*A*	*SA*
Q1	0	0	4	2	0
Q2	2	0	3	1	0
Q3	0	0	2	3	1
Q4	0	0	0	2	4

As showed in Table 2, related to the research question **RQ1**, participants had a slightly higher than neutral agreement with the Q1 statement, related to the adoption of a gamification approach to enhance their motivation (M=3.33, SD=0.51). The gathered qualitative data supports this view as, from the 3 students that answered the open-ended question, a participant reported that *"Gamification would be to support students' motivation"*, another one affirmed that *"I don't know how favorable the competition would be"* and the last one that *"I think it is unnecessary to create competitiveness"*. On its turn, in Q2, participants tended to disagree that the prize would be relevant in such

hypothetical competition (M=2.5, SD=1.22).

Related to **RQ2,** results from Q3 indicate that students usually talked to their colleagues when answering the exercises (M=3.83, SD=0.75). Finally, related to **RQ3,** results from Q4 indicate a highly positive acceptance of the CSPI methodology (M=4.55, SD=0.51).

3.2 Group G2 (with gamification)
Table 3 presents the results from the Likert scale questions related to the G2 Group:

Table 3. N = 8 participants. Statements are presented in Table 1.

	SD	D	N	A	SA
Q2	0	0	3	4	1
Q3	0	1	1	3	3
Q4	0	0	0	1	7
Q5	0	0	0	1	7

As showed in Table 3, related to the research question **RQ1**, participants strongly agreed with the Q5 statement, expressing that the adoption of a gamification approach enhanced their motivation and focus (M=4.87, SD=0.70). In the Q2 statement, participants tended to agree with the importance of prizes in competitions (M=3.75, SD=0.70). Related to **RQ2,** results from Q3 indicate that students usually talked to their colleagues when answering the exercises (M=4.0, SD=0.75).

Finally, related to **RQ3**, almost all students strongly agreed that they liked the CSPI classes and would like the instructor to continue with this approach in the future (M=4.87, SD=0.70). All participants answered the qualitative open-ended question, with all answers being positive towards the adoption of gamification. Answers were classified as the ones that stated that gamification supports students' motivation (4), competition is fun (2), and participants that liked the prizes (2).

3.3 Nonparametric Tests
Overall, as presented in subsections 3.1 and 3.2, considering questions that were administered to both groups (Q2, Q3, and Q4), results from group G2 were always higher than group G1. Also, when comparing the question that assessed whether the gamification would support/supported the student's motivation (respectively Q1 and Q5), results were also higher in the G2 group. These data suggest that gamification was positively accepted and also a factor that potentially supported students' collaboration, motivation, and focus.

Nonetheless, as described in Section 2, specific nonparametric statistical tests would be essential to assess the specifics and whether the results would be or not significant. A two-sided Mann-Whitney U Test (W=1, P < .05) comparing Q1 (G1) and Q5 (G2) answers found, indeed, that although participants in group G1 were skeptical about the potential benefits that gamification could bring to the CSPI classes, participants in group G2 considered the gamification a positive approach to increase students' motivation and focus on the learning process.

However, similar two-sided Mann-Whitney U Test did not find statistically significant differences when comparing G1 and G2 answers for the other questions: Q2 (W=9.5, P=.053); Q3 (W=20, P=.632); and Q4 (W=19, P=.414). This suggests that, although gamification could bring a positive impact on participants' motivation and focus, it does not affect the collaboration between participants, neither the overall acceptance of CSPI classes.

4. Discussion
4.1 Research Questions' Answers
Related to **RQ1**, results suggest that gamification positively affects students' motivation and focus in CSPI classes. Moreover, it was identified that, in regular, not gamified classes, students did not wonder or expect that gamification would bring such benefits. A possible explanation for this could be that the adoption of clickers would be, *per se*, a great novelty. Therefore, in G1, students would still be adapting to the PI approach and clickers technology, not willing, or even considering that gamification could also be included in this scenario. Similarly, in G2, students could not differentiate the benefits of the CSPI approach and gamification, considering both approaches as one unique "whole package". In order to mitigate this, future research could consider a extended period of time, ideally a semester.

On its turn, related to **RQ2,** results suggest that the adoption of gamification does not affect social interaction and collaboration in CSPI classes. This result supports the CSPI original proposal that participants, regardless of gamification be or not be present, should always discuss their answers with peers before submitting them.

The difference observed between the not gamified (G1) and gamified (G2) CSPI version was that the former supported a *cooperative* scenario, in which groups communicated and helped each other, exchanging information and possible solutions. On its turn, the latter supported a *competitive* scenario, with virtually no interaction among different groups.

Future work could investigate the pros and cons of each scenario, including the impact of these approaches on students' learning and individual participation. For example, shy students may feel more comfortable in the *cooperative* environment, interacting only with colleagues in their group. Additionally, they would possibly feel embarrassed to have their responses and performance publicly exposed in the *competitive* scenario. On the other hand, this competition and exposure could force students to focus more on classes, having a positive impact on their learning.

Finally, related to **RQ3**, results (Q4) from G1 and G2 were highly positive and not statistically different, suggesting that students equally appreciated both the non-gamified and gamified CSPI versions.

However, when considering not only the Likert scale related question (Q4) but also the findings already discussed from RQ1 and RQ2 and also the qualitative data, it is fair to consider that, overall, gamification was a positive factor to support CSPI classes.

From the instructor's point of view, however, it is

relevant to consider that the gamified classes required additional efforts, such as managing the scoreboard and also the prizes. This, added to the already required overhead for the CSPI adoption, like the design of questions and content redesign, can be an obstacle for the large-scale adoption of this methodology.

Regarding the prizes, it is natural for participants to feel enthusiastic about winning them (as detected in the G2 group), but its adoption does not seem to be a required feature for a good gamification experience.

4.2 Comparing results to previous works in the area
Klock *et al.* [8] organized a literature survey comparing and evaluating gamification studies through variables like interaction, performance, and user experience. The authors identified that, overall, *"gamification has proved more effective at enhancing student interaction than at increasing engagement and motivation"* [8]. This is, indeed, an opposite result as the findings of this current study (*cf.* Section 4.1).

A possible explanation is that the effects of gamification vary accordingly to the context in which it is used, as each study has its particularities. Therefore, it seems not possible to generalize the gamification outcomes to any context.

Particularly in the CSPI approach, students are encouraged from the beginning to interact with each other. Thus, at least about this factor (interaction), the positive outcomes that could arise from the gamification adoption seem to be limited.

5. Threats to Validity
The main data collected in this study comes from a set of Likert scale and open-ended questions. These questions were mapped to the research questions and then the results analyzed to assess whether variables like motivation, focus and collaboration were reached.

Although this is a typical and literature supported assessment approach [6], it is limited, as the assessment is performed only through questionnaires administered to participants. For example, question Q3 statement relates to assessing whether participants talked to their colleagues when solving exercises. In this case, a determined participant could strongly agree with the statement, but, in practice, had barely communicated with their colleagues. This does not mean that the participant had purposely provided misinformation, but that their perception may eventually be away from reality.

Therefore, it would be interesting to consider and measure more specific and objective data, like the number of interactions between students (from the same group and different ones); the number of correct answers in the CSPI questions; whether students were distracted browsing internet or performing other tasks, etc.

Related to the employed Mann-Whitney U Tests, the study's sample sizes (N1=6 and N2=8) can lead to a minimum *p-value* of .0007 [9], which is inferior to the

adopted level of significance (.05), meaning the results are up to a point valid and reliable.

However, as any statistical test, the greater the number of participants the greater the reliability of the results. For example, the reported Mann-Whitney U Tests result from Q2 had a *p-value of .0536*. Therefore, even the result was slightly above the adopted significance level, it was reported as not statistically significant – what could be different with a greater sample size. In this way, further works could consider larger sample sizes for a more accurate analysis of the data. Also, the participants of this study were mostly instructors from K-12 schools. Future works could be organized in undergraduate CS1 courses.

6. Conclusions
In this work, we present the initial assessment of the gamification adoption into the CSPI. We identified that, although participants originally didn't expect that gamification would affect their experience, overall results did suggest that gamification positively affects participants' motivation and focus, being a positive factor to support CSPI classes. Results also suggest that gamification does not affect social interaction and collaboration. Future works involve the analysis of interaction variables through learning analytics tools.

References
[1] Bonwell, C.C.; J. A. Eison. (1991) Active Learning: Creating Excitement in the Classroom. *ASHEERIC Higher Education Report No. 1*, UGW, Washington, DC.
[2] Zeldbusch, L. *et al.* (2019). SMILE Goes Gaming: Gamification in a Classroom Response System for Academic Teaching. Proc. CSEDU, p. 268-277.
[3] Kelly, M. *et al.* (2015). Response switching and self-efficacy in Peer Instruction classrooms. *Physics Review - ST Phys. Educ. Res.* 11:1.
[4] Hertz, Matthew. (2010). What do "CS1" and "CS2" mean? Investigating differences in the early courses. Proc. of ACM SIGCSE. p 199-203.
[5] iClicker. Available at: https://www.iclicker.com/
[6] Lazar, J. et al. (2017). Research Methods in Human Computer Interaction, Morgan Kaufmann.
[7] Nachar, N. (2008). The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. Tutorials in Quantitative Methods for Psychology. DOI: 10.20982/tqmp.04.1.p013
[8] Klock, A. et al. (2018) Does gamification matter? A systematic mapping about the evaluation of gamification in educational environments. Proc. ACM SAC, p. 2006-2012. DOI: 10.1145/3167132.3167347
[9] Wilcoxon-Mann-Whitney test and a small sample size. Available at: https://bit.ly/3bPVBu0