ILLUMINATIONS

Effect of in-class group clicker-quiz competition on student final exam performance

Jose I. Priego-Quesada,1,2 Irene Jimenez-Perez,1,2 Rosa M. Cibrián Ortiz de Anda,2 Rolando González-Peña,2 and Rosario Salvador Palmer2

1Research Group in Sports Biomechanics, Department of Physical Education and Sports, University of Valencia, Valencia, Spain; and 2Research Group in Medical Physics, Department of Physiology, University of Valencia, Valencia, Spain

Submitted 4 March 2019; accepted in final form 28 June 2019

INTRODUCTION

In the last decades, a university classroom has evolved into an active learning and a more student-centered environment in which the student learns through proactive work rather than as a passive receptor (7, 17). Several advantages of active learning have been observed, such as promoting a more inclusive classroom, improving students’ argumentation, and enhanced student satisfaction and motivation, among others (7, 9, 14). Therefore, making lectures as interactive as possible and creating environments that facilitate learning are important objectives for the medicine and physiology teacher (9, 22).

Gamification is one of the most commonly used active learning strategies. Gamification can be defined as the application of game dynamics to nongame context, with the aim of providing more motivational learning experiences (8, 19). Education is one of the contexts in which gamification has been more widely applied to improve the teaching-learning process (2). It has been observed that the application of gamification in the classroom promotes improvements in cognitive, affective, and social behavior within the classroom (15). In addition, these games usually have a competitive component that increases the students’ level of attention (3).

Among the different options of gamification used in the classrooms, clicker quizzes can provide, when used to enhance the motivational, competitive, and ludic environment, a method of testing to be widely used by the teaching community (5, 20, 23). Testing can be used as summative assessment, diagnostic, or as formative assessment tools (4). As learning and formatting tools, which is the way they were used in the current investigation, clicker quizzes increase retention more than studying alone does, as students receive feedback during testing (4, 21). Therefore, the use of clicker quizzes as a tool to promote learning is an important strategy to investigate (4, 13).

Applications of clicker quizzes are easy to use, as most university students attend lectures using smartphones, laptops, or notepads (10, 11). The advantages of these quizzes are that they provide real-time results, increase student engagement, and make the classroom more interactive (20). Despite these advantages, it is still unknown how these games affect academic performance, given that contradictory and inconsistent results have been observed (5, 16, 18). Furthermore, as mandatory online quizzes have proved to be unpopular with students (1, 5, 12), other alternative strategies need to be assessed (e.g., in-class clicker quizzes). Finally, although there are many studies using online quizzes performed by the students in their free time (5, 12), the number of studies assessing the effect of clicker quizzes conducted in the classroom on academic performance is very limited.

The aim of the present work is to analyze the results in the final score of university physiology students, after the application of two complete in-class sessions of clicker quizzes at the end of the subject. It was hypothesized that this intervention could improve students’ academic performance, given that it may clarify the contents of the subject without involving the negative aspects associated with mandatory online quizzes.

METHODS

Students submitted to the intervention. The intervention was carried out in the academic years 2016–2017, 2017–2018, and 2018–2019 in a subject related to the physiology of the vocal and auditory human organs at the University of Valencia (Spain). The subject is of 4-mo duration and is taught during the first year of university academic training. The subject consisted of 30 lectures (theory lessons) and 10 practical lessons (equivalent to laboratory lessons), and the weekly distribution involved 3 h of lectures and 1.5 h of practical lessons. To analyze the effect of the intervention on the students’ final scores in this subject (corresponding to a final test), the scores from the 3 academic years when the intervention was carried out were compared with the results from the 2 previous academic years: 2014–2015 and 2015–2016. The number of students analyzed in the academic years 2014–2015, 2015–2016, 2016–2017, 2017–2018, and 2018–2019 were 60, 63, 71, 66, and 66, respectively. In our study, these academic years are referred to as cohort 2Y-Before (2 yr before the intervention), cohort 1Y-Before (1 yr before the intervention), Quiz cohort-1 (the first year of the intervention), Quiz cohort-2 (the second year of the intervention), and Quiz cohort-3 (the third year of the intervention). The final test was performed around 3–4 wk after the last quiz session, depending on the official university schedule for each year. The final scores in the subject corresponded to a final test consisting of a 20-question multiple-choice test and 8 short-answer questions, with a maximum score of 10 points and a duration of 90 min. Efforts were made so that the final tests of the 5 academic years evaluated were of similar difficulty: teachers categorized the difficulty of each question and composed a test with 8 questions of little difficulty, 12 questions of moderate difficulty, and 8 questions of great difficulty. The questions included five of the six levels of Bloom’s taxonomy (remember, understand, apply, analyze, and evaluate).
**Intervention.** The intervention was carried out at the end of the academic year during the last two practicals of the subject. The two in-class clicker-quiz sessions were 1.5 h each. One of the sessions focused on the bases of hearing and phonation, and the other on the bases of electrophysiology, thus covering the subject content. It was carried out in practicals as the number of students is lower than in theory classes (~20 students per classroom). The quiz for both sessions consisted of 24 multiple-choice questions with between 2 and 4 possible answers. The questions corresponded to the three first levels of Bloom’s taxonomy (remember, understand, and apply).

For this intervention, Kahoot! software (https://kahoot.com/) was used to develop the clicker quizzes. The Kahoot! is a web application with which the teacher can create test questionnaires for the students to answer during the class using their smart phones, laptops, or notepads (20). Some positive aspects of Kahoot! are that it is a free, fun, and intuitive application (20).

Students were informed to study the subject content before the sessions, to improve their knowledge acquisition. The questions were prepared by the subject professors, taking into account the following aspects:

- They should only include contents and key concepts that had to do with the learning objectives developed in the teaching guide of the subject (example of a learning objective: To relate the phonation process with the subjective magnitudes of sound).
- Each question was posed in such a way that it could be answered clearly within 30 s, which was the maximum time stipulated for the application [example of quiz question: The sound generated next to the vocal cords is composed: a) only by the fundamental harmonic; b) only by the formants; c) by the fundamental frequency and its harmonics; d) by the fundamental harmonic and the formants].
- In no case should questions be similar to those that would appear in the final exam of the subject (example of final exam question: Relate the physical magnitudes of the emitted sound with the subjective magnitudes of the perceived sound).

During the session, the following actions were performed:

- It was undertaken as group work, not individually, so as to avoid individual winners or losers, and a pseudonym was used for each group.
- The classroom was divided into five groups of four students each. The choice of five groups was because only five participants are visualized in the application’s “visualization of results.” By dividing the class into five groups, all groups could, therefore, see their evolution. Groups were self-selected by students to promote groups with the highest level of cohesion possible. The teacher encouraged the participation of all members, by obliging each of them to take part during the session in explaining the reasons for a particular answer.
- The protocol for each question was as follows. 1) Sufficient time for the question to be displayed and discussed by the group before choosing their answer (Fig. 1). 2) Display of the right answer and wrong answers. Time is provided for the teacher to clarify why each answer was right or wrong. If the teacher detected a deficit on the part of the students in some concept or content, a more detailed explanation was provided. 3) Display of the results. The application displays the ranking of the group for each question, taking into account whether the answer was right, as well as the time used in answering. After this ranking per question, it also shows the ranking of the total number of questions answered so far. Competition time is encouraged to increase student motivation. Humor is also used as a motivation tool (e.g., joking about the ranking or some of the wrong response options), and, for failing groups, not to feel embarrassed and understand that the purpose of the session was to improve their learning process.

**Statistical analysis.** To analyze the effectiveness of the intervention implemented, the results of the students’ final scores from the Quiz cohort-1, Quiz cohort-2, and Quiz cohort-3 academic years were compared with the results of the 2 previous academic years in which the clicker-quiz sessions were not carried out (cohorts 2Y-Before and 1Y-Before).

Data were analyzed using SPSS Statistics 21.0 (IBM, Armonk, NY). The normality of the final score distribution for the subject over the 4 academic years was confirmed by the Kolmogorov-Smirnov test.
and homogeneity by Levene’s test ($P > 0.05$). The score over the 5 yr was then assessed using a one-way ANOVA with Bonferroni post hoc comparisons. For the pairwise comparisons, Cohen’s effect size (ES) was computed and was classified as small (ES 0.2–0.5), moderate (ES 0.5–0.8), or large (ES>0.8) (6). Finally, the students’ final scores were categorized into different qualification grades: D (score lower than 5 out of 10), C (score between 5 and 6.9 out of 10), B (score between 7 and 8.9 out of 10), and A (score equal to or higher than 9 out of 10). The distribution of each qualification grade over the 4 academic years was assessed using a chi-squared test. Data are reported as means ± SD, with 95% confidence intervals of the differences between academic years (95% CI). Statistical significance was defined as $P < 0.05$. 

**RESULTS**

The purpose of this investigation was to analyze the academic performance of university physiology students who had undergone two sessions of clicker quizzes using the Kahoot! Application (Fig. 1) at the end of a physiology course during 3 academic years (cohorts Quiz-1, Quiz-2, and Quiz-3), with those of the 2 previous years who had not undertaken those clicker quizzes (cohorts 1Y-Before and 2Y-Before). The two sessions occurred in class as a form of group-clicker quiz application (cohorts Quiz-1, Quiz-2, and Quiz-3) or not using it (cohorts 2 yr before (2Y-Before) and 1 yr before (1Y-Before)). Differences between courses: *$P < 0.05$ and ***$P < 0.001$.

Figure 2 shows the final score in the subject of the 5 academic years assessed (the first 2 yr without the intervention and the last 3 yr when the clicker-quiz sessions were performed). The final evaluation scores were higher with the quiz session application than without (Fig. 2). The first academic year using the quiz session application presented higher academic performance than the 2 previous academic years without the intervention (95% CI of the difference with 2Y-Before: 0.0–1.3 points, $P = 0.03$ and ES = 0.5; 95% CI of the difference with 1Y-Before: 0.3–1.3 points, $P < 0.001$ and ES = 0.7). The second academic year using the quiz session application presented higher academic performance than in 1Y-Before (95% CI: 0.0–1.3 points, $P = 0.04$ and ES = 0.5). The third academic year using the quiz session application presented higher academic performance than the 2 previous academic years without the intervention (95% CI of the difference with 2Y-Before: 0.0–1.4 points, $P = 0.03$ and ES = 0.5; 95% CI of the difference with 1Y-Before: 0.3–1.6 points, $P < 0.001$ and ES = 0.8).

Figure 3 shows the percentage of each qualification grade during the five courses assessed. The qualification grade was categorized as follow: D (score lower than 5 out of 10), C (score between 5 and 6.9 out of 10), B (score between 7 and 8.9 out of 10), and A (score equal to or higher than 9 out of 10). The chi-squared test showed that the distribution of the qualification grade differed over the academic years ($P < 0.001$). No clear effect can be observed in the percentage of students with the higher grade (A) or in the percentage of failed students (D). However, a different distribution of percentages of students with C and B scores can be observed with the application of clicker-quiz sessions. Students with B scores increased by ~20% as a result of the decrease in those with a C score.

**DISCUSSION**

This study investigated the impact on academic performance of introducing two sessions of clicker quizzes at the end of the course. The 3 academic years with the intervention were then compared with the 2 previous academic years without it. The main result was that academic performance moderately increased with the intervention. More specifically, although the percentage of failed students did not alter, the distribution of the scores of the students who passed changed, increasing the students with a B score by ~20% with the corresponding decrease of students with a C score.
Various studies have been carried out over the last decade aimed at exploring the effect of online quizzes on academic performance and knowledge acquisition. Kibble (12) administered optional online quizzes in medical physiology 2–4 wk before summative exams, observing an improvement in those students who had taken these online quizzes. Lee and colleagues (16) observed that, although online quizzes enhanced social presence and positive perception of learning compared with traditional lectures, no differences were observed in learning outcomes. Brown et al. (5) administered mandatory online quizzes the day before an in-class chapter test in an anatomy and physiology subject, observing that, for some tests, performance improved, but not for all. However, it is important to mention that performance was not impaired (5). In contrast, Shaikh and colleagues (23) used an online quiz reinforcement system. This system consisted of an adaptive platform that sent questions repeatedly in time to the participants to facilitate memorization. The frequency of question repetition depended on the respondents’ success (13 days if they answer correctly and 7 days if not). Knowledge acquisition was observed using this system (23). Our study observed an increase in the final test score for the subject after applying the in-class quiz sessions (Fig. 2). Our intervention was performed 3–4 wk before the final test, which may condition the results. Although less time there is between the quiz and the final test could improve results, this time could not be changed in our institution because the quiz sessions were performed in the last weeks of the classroom calendar, and the date of the final test in the exam period is not modifiable. In addition, it is important to mention that the magnitude of the effect of the increase in test score was moderate, and the second year of application only presented differences with one of the previous years. It is, therefore, understandable that previous studies found no effect of the online quizzes on academic performance (5, 16). The analysis of the different qualification grades could explain more clearly the effect of the intervention than the numerical final test score (Fig. 3).

The results of our study suggest that the application of two in-class sessions of clicker quizzes did not affect the percentage of students who are excellent (A score) or the students who fail (D score). For the A score group, it can be speculated that there are ~10% of hard working, intelligent students who value these interventions, even though they do not lead to a higher score. Similarly, for the D score group, it can be speculated that there are ~10% who went to the session without having studied, do not participate, and leave without having learned very much. Their qualification, therefore, remains much the same. However, it can be observed that for most students, who are in the center of the distribution, it does have an effect by increasing the percentage of B scores and reducing the C scores. We believe that these results show that the in-class clicker-quizzes sessions helped to clarify contents and consolidate knowledge. As a qualitative comment, the teachers perceived that, in the second quiz session, students prepared the content more deeply in advance, probably as a result of their experience of competition in the first session. Therefore, this intervention was perceived by the teachers to be a strategy that can motivate students to study more in advance.

Most of the previous studies discussed above were online formative quizzes performed by the students outside the classroom. Those procedures were perceived as unpopular by the students, mainly due to time constraints (e.g., higher work load and difficulty to combine with personal life) and the lack of feedback, among others factors (1). Our procedure, therefore, had the advantages over the previous ones, of being in class and during course time, and providing concise and immediate feedback on correct and incorrect answers (1, 18).

To provide a credit score for performing the clicker quizzes is also an issue that needs investigation. Marden et al. (18) assessed the effect on academic performance of four different interventions with clicker quizzes: 1) 5% credit, unsupervised and with no use of textbooks and notes; 2) 5% credit, supervised and with no use of textbooks and notes; 3) 7.5% credit, supervised and with the use of textbooks and notes; and 4) 2% credit for excellent scores, unsupervised, with the use of textbooks and notes and with the same questions and multiple choices in the final evaluation of the physiology course. Although they observed that only the fourth intervention obtained a higher academic performance, they were not sure if learning was enhanced or students memorized the questions and answers of the quizzes (18). In our intervention, attendance at the sessions was obligatory, no credit was provided for participants and questions and answers were not the same as in the final test. No credit guaranteed a good atmosphere in the classroom, and the improvement in the distribution of academic performance was attributable mainly to the feedback and clarifications provided by the teacher.

Kahoot! application was used in this intervention as a clicker-quiz software. Although this application has some advantages, such as it is free and answers are automatically collected and scored immediately (20), the restricted number of characters accepted is a drawback that needs to be taken into consideration. Questions and answers have a character limit of 95 and 60 characters, respectively. This could make it difficult to formulate the teacher’s questions as well as the level of knowledge (Bloom’s taxonomy) of the questions. In this sense, the questions of the Kahoot! had a lower level of knowledge (reaching the third level, “apply”) than those of the final exam (reaching the fifth level, “evaluate”). For this reason, it is important that teachers understand that the application for which the Kahoot! was used has its limitations as far as the high levels of knowledge are concerned, and, therefore, its utility is greater for the lower levels of knowledge.

Future studies on our type of intervention could address whether there are differences in academic performance between students who studied before the quiz sessions and students who did not.

Conclusion. The application of in-class clicker quizzes at the end of a physiology course (language and hearing organs), aimed at clarifying, checking, and retaining the course content, was beneficial, improving the average score of the classroom and, more specifically, increasing the percentage of students with a B score (7–8.9/10 points) by 20% with the consequent decrease in those with a C score (5–6.9/10 points).

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Advances in Physiology Education • doi:10.1152/advan.00032.2019 • http://advan.physiology.org
Downloaded from www.physiology.org/journal/advances (052.011.211.149) on January 13, 2020.
REFERENCES


