Journal of Effective Teaching in Higher Education

The Effects of Gaming on University Student Quiz Performance

Natalie R. Andzik, Northern Illinois University, Nandzik@niu.edu
Corinne M. Gist, Elle E. Smith, Menglin Xu, Nancy A. Neef, The Ohio State University

Abstract. The use of competitive games to increase classroom engagement has become common practice among many teachers and university professors. However, it is unclear if using games, as an assessment tool, is a viable way to increase student performance. This study examined the effects of administering point-earning quizzes through a game-based system, Kahoot!, visible to all in the classroom versus privately on an electronic device. The quiz scores of 56 undergraduate students, enrolled in one of two special education courses, were evaluated. A linear regression was used to compare student scores across the two conditions as well as performance over the course of a 15-week semester. No significant difference in quiz scores was found between the two conditions, and quiz scores in both conditions improved similarly over time. Sixty-eight percent of the students reported preferring to take the quiz privately on an electric device as opposed to on Kahoot! Limitations and recommendations for practitioners are discussed.

Keywords: Higher education, college instruction, game-based assessment, student-response system

Instructors in higher education seek to engage students in the learning process while simultaneously assessing their skills and knowledge. Formative and summative assessments are commonly used in the college classrooms for this purpose (Brookhart, 2004). Summative, or formal, assessments include systematic ways of assessing student learning and providing instructors with information that is useful in making final decisions, such as grade assignments (Brookhart, 2004). These assessments often include incentives to perform well (e.g., points) associated with them and may include quizzes, midterms, and final exams. Informal, or formative, assessments refer to a wide variety of methods that teachers use for guiding student learning, instructional change, and other improvements (Brookhart, 2004). Examples of formative assessments include active responding, think-pair-share, and choral responding. A popular version of active responding, often used in the college classroom, includes classroom response systems (CRS), also called student response systems (SRS) (Wang, Zhu, & Sætre, 2016). SRS may employ individual remotes that students use to respond to questions that are projected on the screen and a handheld transmitter that collects and displays the student's responses. Other classroom response technologies are web-based and allow students to use cell phones or other personal mobile devices to answer questions (Wang, 2015).

SRS allow instructors to instantaneously collect data from student responses and display the answers on a classroom projection screen where both students and

instructor can see and discuss them (Wang, 2015). Researchers have reported many benefits to using SRS in the classroom including improved attendance, increased focus among students, improved student engagement, increased learning performance, improved teaching, and improved interaction between teacher and students (see Kay & LeSage, 2009, for a review). A newer SRS system that has become increasing popular is Kahoot! Kahoot! is an example of game-based student response systems (GSRS) that combine the benefits of SRS with those of game-based learning. As a game-based system, Kahoot! is intended to transform the classroom into a game show, where the students are the competitors and the instructor is the host (Wang, Zhu, & Sætre, 2016). Kahoot! gamifies the student response process by using a graphic user-interface, special audio, and competition. These effects are designed to attract and engage the students.

Game-based learning has been received positively when used in the university classrooms. For example, in a recent study, students reported that classroom-learning games increased their enjoyment of the learning process (Crocco, Offenholley, & Hernadez, 2016; Robinson, 2014). In addition, two experimental studies found that higher quiz and test grades were achieved when game-based learning was used as a study strategy (Neef, Perrin, Haberlin & Rodriques, 2011; Robinson, 2014). Despite the benefits, prior to considering game-based learning in classrooms and deciding on a medium for its delivery, practitioners should consider the competitive aspect of this learning strategy as well as the makeup of the students in the classroom. The element of competition can aid in student learning as evidenced by a study that used competitive game-based learning as a way of demonstrating learning and self-reported motivation among undergraduate students (Cagiltay, Ozcelik, & Ozcelik, 2015). However, Neef and colleagues (2011) used a different, cooperative approach with gaming, and although there was not a noticeable difference between the gaming and control conditions, the groups expressed they "felt bad" for the other team and expressed they wished them to do well on the quizzes.

One obvious benefit when incorporating GSRS in the classroom is the natural increase in participation among students. University instructors and professors are often up against distracting technology (e.g., cell phones, Internet) and likely decreased participation when teaching larger groups of students (Rocca, 2010). Public posting through a "score board" can be beneficial when capturing larger audiences of students and inadvertently pressures students to prepare in advance to avoid any possible embarrassing effects of placing last in comparison to their peers in a game. Anecdotal reports suggest other benefits of GSRS when used as informal assessments or study tools (Plump & LaRosa, 2017; Wang, 2015). Although few experimental studies have been conducted, it is reasonable to assume that GSRS could be used as a formal assessment tool.

Currently, only one study has looked at the effects of Kahoot! in this way. Wang, Zhu, and Sætre (2016) compared the effects of using Kahoot! SRS (i.e. Clicker) and paper tests as a formal assessment. Pre-and post-tests were used to assess the learning outcomes of the lectures and a questionnaire was given to obtain feedback on the students' engagement, enjoyment, concentration, and motivation.

Students who took the Kahoot! quiz reported higher levels of motivation, enjoyment, engagement, and concentration during the assessment than those who were assessed using paper tests or SRS. To assess the learning outcomes, quiz scores were analyzed using the Mann-Whitney test. No significant difference was found between the quizzes taken on Kahoot! and the quizzes taken on paper. The SRS condition was not included in the learning outcome analysis.

One limitation of the Wang, Zhu, and Sætre (2016) study included only exposing the participants to one condition. Therefore, the students who took the quiz using Kahoot! could not compare the gamified approach to the SRS or paper assessments. In addition, data were collected on only one quiz. Finally, the researchers did not use any incentive (e.g., points, rewards) for the completion of the quiz. Incentives have been shown to increase, even if temporarily, the motivation of students (Bartel & Hagel, 2014)

The current study aimed to extend the work of Wang, Zhu, and Sætre (2016) by increasing the number of assessments, having participants take the quizzes in both conditions, and including incentive (i.e., points in the class). This allowed the participants to compare their experiences in each condition and created a larger sample of quiz scores. The aims of this study were to (1) determine if there was a difference in scores when quizzes were taken on Kahoot! versus privately on an electronic device, (2) determine if student performance changed over time, and (3) determine if students preferred one quiz method over the other.

Methods

Participants and Setting

Fifty-six undergraduate students from a large Midwestern university were included in this study. After obtaining University Institutional Board Approval, students were recruited from either Introduction to the Special Education Profession (n=21) or Applied Behavior Analysis (ABA) for Teachers (n=35), both instructed by the second author. Data were collected during weekly class sessions (2 h 45 min) over a 15-week semester. All sessions occurred on campus and were conducted in university classrooms.

ABA for Teachers was designed to provide an overview of applied behavior analysis, and the aims of the course were to introduce students to the basic principles by which humans learn social, academic, physical, and other skills. These basic principles include, but are not limited to, positive and negative reinforcement, extinction, punishment, discrimination, and stimulus control. Pre-service teachers took this course in their first semester after being accepted into the licensure program (typically their third year in college). Introduction to the Special Education Profession was a course that provided an overview of the education of exceptional learners. Students learned about the various categories of special education, including the criteria for each category, prevalence, demographics, and typical interventions. In addition, students studied special education laws as well as special education as a professional discipline, including its history, current issues and

challenges, and contemporary research-based instructional practices. This course was designed for pre-service teachers to take before their acceptance into the licensure program (often taken during their second year in college).

Both courses were introductory in nature and chapter quizzes were designed to (a) ensure students read the material prior to lecture, (b) assess the level of comprehension of the text of all students, (c) prompt students to critically think about the content prior to lecture and come prepared with questions. The introductory level of these courses highlights the primary level of learning—remembering and recalling facts, initial acquisition of new concepts, and understanding of the materials and how they affect learners with disabilities. Each quiz was worth 10 points and accounted for 30% (Introduction to Special Education) and 35% (ABA for Teachers) of the total grade in the course.

Procedure and Materials

An alternating treatment design was used in which the participants shifted between the two quiz conditions (Johnston & Pennypacker, 2009). Students were randomly assigned to one of two groups (group A and group B). To determine if there were any outliers in the group, a pre-test was conducted to establish a baseline measure of the participants' basic knowledge of the subject matter. No outliers were found. The groups were exposed to one of two conditions each week, private and game-based. At the end of each class session, students were notified which quiz condition their group would be in the following week. At the beginning of each class, the instructor administered the quizzes, which included 10 multiple-choice questions based on the week's assigned reading. Each quiz had a time limit of 20 minutes and students could not change their answers once the selection was made. During the private condition, students took the quiz on a computer or personal electronic device (e.g., iPad, smart phone, laptop computer) in a nearby conference room supervised by either a graduate student or the classroom instructor. Only the students taking the quiz could see their screen and answers.

The group of students in the game-based condition took the same quiz as the group in the private condition, but these students used Kahoot! as a group in the classroom. Kahoot! is a free online platform used by teachers to assess the knowledge of their students. The quiz questions were displayed one by one, each for 2 minutes, or until all students had answered. If all students answered before the 2-minute time cap, the question automatically left the screen. The questions were projected on a large screen along with four possible responses. The students answered by choosing the corresponding symbol on the board that matched that on their personal electronic devices. After every question, the correct answer, a bar graph with the distribution of how the students answered, and the top three players were displayed on the main screen. Students received individual feedback on their personal devices including the number of points received, their ranking compared to the rest of the group, how far the student was behind the student ranked above them, and the correct answer. At the end of a game-based session, the top three-point earners' names were displayed on the classroom screen. Points were earned

for answering quickly and correctly. The faster a student answered correctly, the more points they received.

Students in both conditions could retrieve answers to the quiz questions on the classroom website after the class period had ended. There was no data collected regarding which, if any, of the participants checked the answers online following the quizzes. Each group participated in the private and public condition an average 6.5 times (range 6–8). Although each quiz was created to be equally challenging, it was not possible to compare difficulty of the material across quizzes; therefore, analysis focused on within-unit comparison of quiz scores for private and public groups. Although proctoring for each testing condition varied between an instructor or graduate student, the students' experience in each testing condition were identical.

Interobserver Agreement

Interobserver agreement (IOA) was assessed on all quizzes across all sessions. All quizzes were written and scored by the first, second, and third authors. The private group quiz scores were graded by a university online platform (i.e., Canvas), and Kahoot! scores were automatically graded by the website. Each week, all scores were double-checked for accuracy by the first, second, and third authors and 100% agreement was found.

Social Validity

Social validity was assessed through an anonymous questionnaire given at the conclusion of the study. One participant opted not to fill out the survey. Participants were asked 11 questions, including 7 multiple-choice and 4 open response questions, and were provided space for additional comments. These questions were designed to assess the participants' satisfaction with the procedures of the study (Wolf, 1978).

Statistical Analysis

First, the quiz scores were re-scaled to fall into the range of 0 and 1, so that the scores collected from 12 quizzes were comparable. Second, data were transformed from the original wide format to long format in order to fit the regression model. Specifically, the variable of Type (0, 1) was created to represent the private condition and the Kahoot! condition respectively, the variable of Time was created to indicate the week that the quiz was taken, and the variable of Quiz Scores (numeric score) was added to serve as dependent variable. The long format data contains 672 quiz scores (i.e., 56 students at 12 time points). Third, descriptive statistics were run via SPSS 24.0 to produce mean and standard deviation (SD) for the private condition and Kahoot! condition. Finally, a linear regression model was fit using SPSS with the quiz outcome serving as the dependent variable and Time and Type serving as the independent variables.

Results

Thirty-nine of the 672 quiz scores were missing (rate = 5.8%); therefore, listwise deletion was adopted in the subsequent analyses due to the low missing rate. The mean scaled quiz outcome was .725 (SD = 0.183) for the private condition (*n*=337) and .707 (SD = 0.193) for the Kahoot! condition (*n*=296), revealing no pronounced difference across quiz types. Table 1 displays the results of the regression model. Type had no significant effect on quiz outcome, β = -.018, SE = 0.015, p = .216, which indicated that there is no prominent difference between the private condition and Kahoot! condition after controlling for time. However, a significant time effect on quiz performance was detected, β = .006, SE = 0.002, p = .003, which indicated that students experienced significant growth over the semester, regardless of quiz type. Students quiz scores improved as the semester went on. Type and Time explained 1.6% of variance in quiz outcome. Figure 1 displays the growth pattern of Quiz performance by Type. Students progressed at similar speeds for both conditions.

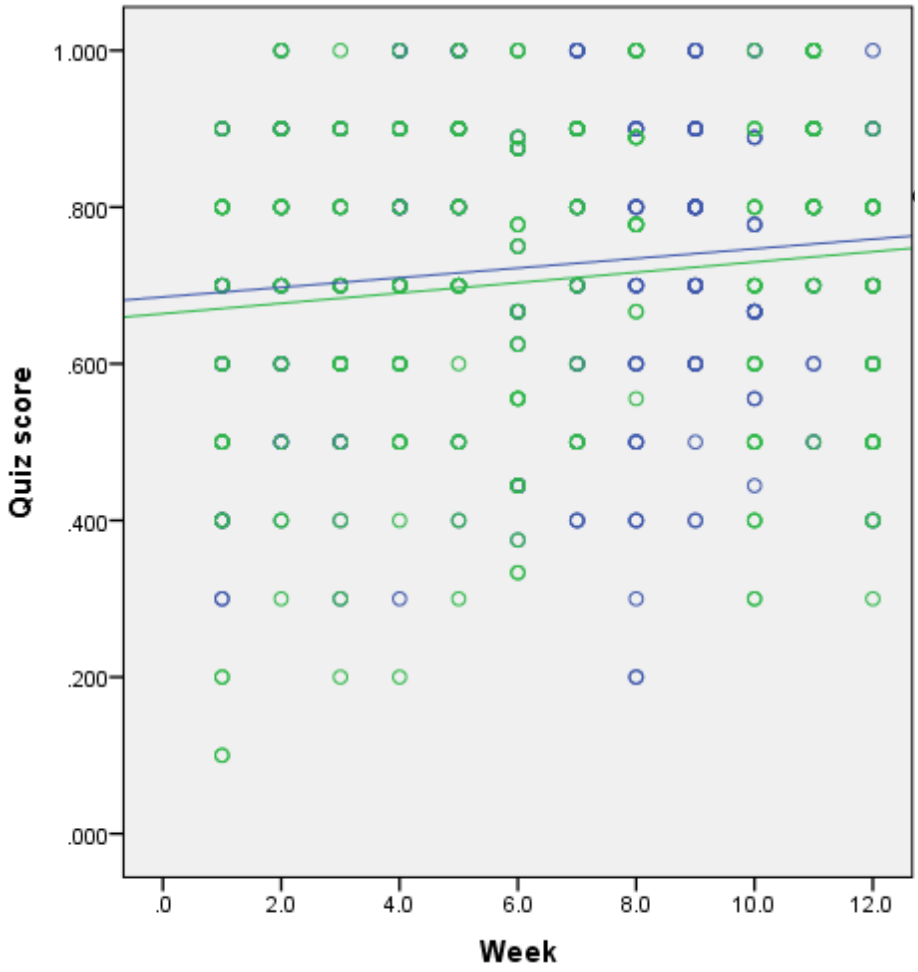

Figure 1. Quiz growth outcome by Type.
Note. Blue line: Private; Green line: Kahoot!

Table 1
Results of Regression Model

	β	SE	t	p
Intercept	0.684	0.017	39.833	<.001
Type	-0.018	0.015	-1.239	.216
Time	0.006	0.002	2.976	.003
R^2 = .016				

Note. Type: 0: Private condition; 1: Kahoot! condition; Time: the week that the quiz was taken.

Social Validity

Table 2 displays the results of the questionnaire completed by 55 participants. When asked which assessment type the participants scored higher on, most indicated they did well on "Both" (*n*=24) or that they did better on "Private" (*n*=23). In response to which assessment type the participants studied for more, 95% reported "Both" (*n*=52). When asked which quiz type increased their content knowledge, most participants reported "Private" or "Both" (*n*=20 for both responses). The majority (65%) of the participants reported they enjoyed the Private condition over the Kahoot! condition (n=38). Most participants (73%) would prefer to take an assessment on a private device versus on Kahoot! in the future. When asked if there was anything the participants disliked about the assessment conditions, the majority (89%) of the participant reported "yes" (*n*=49) for a dislike about Kahoot! and 47% (*n*=26) reported "yes" for a dislike about Private. The students provide additional information on what they liked and didn't like about each condition as described below.

The researchers coded the open-ended questions based on common themes across the answers. Given that the participants could provide multiple answers for each question, the *n* score reflects the number of responses, not the total number of participants. The participants were asked to evaluate the components they liked for both assessment types. When asked about Kahoot!, the majority (*n*=27) of the responses commented positively about the immediate feedback following each question. Some of the participants reported that the points and scoring features of Kahoot! made the assessment fun and motivating (*n*=10). Fourteen participants reported that there was nothing they disliked about Kahoot! The participants' evaluations of the private condition highlighted that many participants (*n*=29) liked that the assessment was self-paced and there was not a per question time limit. Additional positive responses included the non-distracting format, the similarity to traditional assessments (*n*=14), private scores (*n*=8), and the lack of pressure for the participants to respond quickly (*n*=7).

Table 2
Social Validity Questionnaire Responses

Question	Answers		
	Kahoot!	Private	Both
On which quiz did you receive a higher score?	16%	42%	44%
Which quiz type did you study for more (duration, intensity, etc.)?	2%	4%	95%
I feel my content knowledge about the chapter was stronger after taking the _____ quiz.	27%	36%	36%
Which quiz did you enjoy more?	25%	65%	4%
In future classes, I would prefer to take quizzes on a _____format.	22%	73%	N/A
	Answers		
	Yes	No	
Did you dislike anything about the Kahoot quizzes?	*n*=49 (89%)	*n*=6 (11%)	
Did you dislike anything about the Private quizzes?	*n*=26 (47%)	*n*=25 (45%)	

Participants were asked to make suggestions for improving the assessments. The majority of comments about Kahoot! recommended turning off or reducing the volume of the music during the assessment. Many participants found the music distracting (*n*=24). Additional recommendations including eliminating the time requirement for answers (*n*=8), being able to change answer selections (*n*=8), and removing the display of the number of people who answered incorrectly (*n*=6). Several participants recommended using Kahoot! for review instead of for a grade. The most commonly provided recommendation for the Private condition was related to the format of the assessment. Specifically, participants wanted to change answers on previously completed questions (*n*=21) and to skip and/or see all the questions at once (*n*=7).

Discussion

Similar to the findings of Wang, Zhu, and Sætre (2016), there was no significant difference between private quiz performance and Kahoot! quiz performance. Students in both groups showed growth in quiz scores over the course of the semester, but these results were not related to the quiz condition. Finally, students reported preparing for both testing conditions in the same way and reported a preference for the Private condition over Kahoot! This study extended the work of Wang, Zhu, and Sætre (2016) by exposing students to both conditions multiple times.

Limitations

When considering the results of this study, a few limitations should be considered. First, future researchers should consider using a larger sample with random sampling when comparing multiple methods for improving student performance on assessments. In addition, Kahoot! carries some limitations as an assessment tool and should be used with caution. Some students with disabilities could not participate in this study, as the nature of the group assessment does not allow for extended time or access to a distraction-free environment. The individuals with disabilities enrolled in these courses were offered pencil/paper tests to accommodate their needs and thus their scores were not included in the analysis. Second, Kahoot! audio was reportedly distracting to a lot of students as many requested the music be turned down, and several students muted the audio feedback on their personal devices. Third, some students who are English Language Learners could not participate as they required more time to process the language presented on the screen and needed pauses for frequent clarification.

Implications for Practitioners

Practitioners who adopt GRSS in their classroom should take the findings from this study into consideration. Participants overwhelmingly preferred the assessments conducted in the private setting. The elements of Kahoot! that are designed to be stimulating and enticing (e.g., music, feedback) were reported to be distracting to students and were not supportive of a typical quiz-taking environment. However, one important benefit to Kahoot! was the immediate feedback students received during the assessment. Students appreciated this immediacy of the answers provided when using Kahoot! Students assessed using the Private condition were unable to access the answers to the quiz questions until after class. Practitioners should consider providing immediate feedback during assessment whenever feasible. Immediate feedback on a web-based platform could include availability of answers following the completion of a quiz, availability of answers following each question, or availability of answers after all students have completed a quiz. An additional option that does not include a web-based platform might have the instructor verbally reviewing the answers after an in-person assessment.

In addition to immediate feedback, many participants indicated a desire to be able to go back to previous questions on the quiz. Some students commented that they remembered the answer for a previous question after they submitted their answer or that they would have liked to skip and then return to a question. Of course, if immediate feedback were given directly after the submission of an answer, backtracking would not be possible. Practitioners might poll their students at the onset of the course to see where their preferences lie.

An additional benefit of using Kahoot! is the built-in data collection system. After each quiz is administered, a report of the results is generated and available for the instructor to download. The report includes specific information regarding how each student answered individual questions as well as percentage correct information for

each question. These reports can provide the instructor with immediate information regarding the students' comprehension of the content and highlight what areas may need to be re-taught or reviewed. This feature is a benefit whether Kahoot! is being used as a formal or informal assessment.

Conclusion

GSRS provide a fun and innovative way of getting students to be excited about learning in the classroom. The addition of games in the classroom can be refreshing to students in university and K-12 schools. However, when using a public, game-based platform for formal assessment, students reported the format to be less preferable when compared to a private, more traditional platform. The findings from this study may indicate that Kahoot! is best used to informally assess student learning and not as a formal method of assessment. Practitioners should use Kahoot! as a strategy for formally assessing their students with caution; however, game-based learning has its benefits and should be considered when reviewing content, preparing students for an upcoming exam, or other non-weight bearing activities to supplement instruction. Kahoot! may also be most beneficial when used as a baseline assessment, a presentation tool to keep students engaged and check for understanding, or as a review or study tool for formal assessments. When developing formal assessments, instructors should consider learners with special needs and should also consider immediacy of feedback and allow for students to move freely between quiz questions during an assessment.

References

Bartel, A., & Hagel, G. (2014). Engaging students with a mobile game-based learning system in university education. Engineering education towards openness and sustainability: *Proceedings of the IEEE Global Engineering Education Conference (EDUCON)* (pp. 957–960). Istanbul: IEEE. doi:10.1109/EDUCON.2014.6826215

Brookhart, S.M. (2004). Assessment theory for college classrooms. *New Directions for Teaching and Learning, 100,* 5–14. doi: 10.1002/tl.165

Cagiltay, N. E., Ozcelik, E., & Ozcelik, N. S. (2015). The effect of competition on learning in games. *Computers & Education*, *87*, 35–41.

Crocco, F., Offenholley, K., and Hernadez, C. (2016). A proof-of-concept study of game-based learning in higher education. *Simulation & Gaming*, *4,* 403–422. doi: 10.1177/1046878116632484

Johnston, J.M., & Pennypacker, H.S. (2009) *Strategies and tactics of behavioral research* (3rd ed.). New York, NY: Routledge.

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, *53*, 819–827. doi: 10.1016/j.compedu.2009.05.001

Neef, N.A., Perrin, J.A., Haberlin, A.T., & Rodriques, L.C. (2011). Studying as fun and games: Effects on college students' quiz performance. *Journal of Applied Behavior Analysis, 44*, 897–901. doi: 10.1901/jaba.2011.44-897

Plump, C. M., & LaRosa, J. (2017). Using Kahoot! in the classroom to create engagement and active learning: A game-based technology solution for

eLearning novices. *Management Teaching Review*, *2*, 151–158. doi:10.1177/2379298116689783

Rocca, K. A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education*, *59*, 185–213.

Robinson, F. F (2014). It's a game! Evaluation of a classroom game to enhance learning in an introductory counseling course. *Innovative Teaching*, *3*, 10. doi:10.2466/o7.08.IT.3.10

Wang, A.I. (2015) The wear out effect of a game-based student response system. *Computers and Education*, *82*, 217–227. doi: 10.1016/j.compedu.2014.11.004

Wang, A. I., Zhu, M., & Sætre, R. (2016). *The effect of digitizing and gamifying quizzing in classrooms*. In T. Connolly & L. Boyle (Eds.), *Proceedings of the 10th European Conference on Games Based Learning* (pp. 729–737). Paisley, Scotland: Academic Conferences and Publishing Internatioal.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*, 203–214. doi: 10.1901/jaba.1978.11-203