



Comparing the effectiveness of badges and leaderboards on academic performance and motivation of students in fully versus partially gamified online physics classes

Sebiha Balci¹ · Jonathan M. Secaur² · Bradley J. Morris¹

Received: 15 December 2021 / Accepted: 28 February 2022 / Published online: 12 March 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Gamification, or the intentional use of gaming elements in non-game contexts, has been touted as a promising tool to improve educational outcomes in online education, yet the evidence regarding why it might work and its effectiveness is inconclusive. One reason is that previous research has often included several gamification tools together, neglecting that each gamification tool can vary in effectiveness. In order to evaluate their relative impact, two frequently used gamification tools, badges (i.e., digital credentials given for achievements) and leaderboards (i.e., digital rankings based on performance), were compared for their effectiveness on the academic performance and motivation of students. Two experiments were conducted in two online undergraduate physics courses taught via a learning management system. In Experiment 1 (N=102), badges and leaderboards were implemented in only one part of the course grading system (i.e., quizzes). In Experiment 2 (N=88), all course grading system was gamified (i.e., quizzes and assignments). Four groups were created by random assignment of participants: badges-only, leaderboards-only, badges with leaderboards, and control (i.e., no badges, no leaderboards). Academic performance was measured by comparing quiz scores among groups in Experiment 1 and both quiz and assignment scores in Experiment 2. Participants filled out a self-report motivation survey about badges and leaderboards at the end of the study. Two experiments yielded similar results: badges and leaderboards did not affect participants' academic performance; however, most students approached them positively as motivational tools and wanted to see them in future online classes.

Keywords Badges · Leaderboards · Gamification · Online learning · Motivation

✉ Sebiha Balci
sbalci@kent.edu

¹ Department of Lifespan Development and Educational Science, Kent State University, 405 White Hall, Kent, OH 44242, USA

² Department of Physics, Kent State University, 103 Smith Hall, Kent, OH 44242, USA

1 Introduction

Online education has become a significant aspect of K-12 and higher education (Means et al., 2014), with almost 33.1% of university students taking at least one distance education class in the fall of 2017 (Grinder et al., 2019). However, the COVID-19 global pandemic accelerated the shift to online education in an unpredictable way, such that nearly 96% of USA colleges in the Fall 2020 semester operated fully or primarily online (Dennon, 2021). Despite the numerous opportunities in online education, serious challenges are yet to be overcome. One issue is how to motivate students when they are responsible for their own learning experience without the support afforded by face-to-face interaction with instructors (Dennen & Bonk, 2007; Mahle, 2011). As college students had to take most of their classes online during the COVID-19 pandemic, the challenges of online education, such as staying motivated, engaging with the course content, and participating in the class activities, became more prominent due to social isolation (Nair, 2021; Zainuddin et al., 2021).

One potential tool to provide motivational support to online college students is gamification, or the intentional use of game design elements, such as badges, leaderboards, points, trophies, narrations etc., in non-game contexts (Deterting et al., 2011; Faiella & Ricciardi, 2015; Seaborn & Fels, 2015). Incorporating game features into courses could appeal to online students because games are inherently engaging and motivating for users (Dichev & Dicheva, 2017; Hamari, 2017; Seaborn & Fels, 2015). Badges and leaderboards are the most frequently implemented game design elements (hereafter gamification tool) (Hamari et al., 2014; Seaborn & Fels, 2015). Badges are defined as digital credentials awarded for acknowledging achievements and skills, and leaderboards are digital rankings created based on performances (Alaswad & Nadolny, 2015; Grant, 2014). However, as many of the gamification studies have implemented more than one gamification tool simultaneously, it is challenging to identify the effects of each tool (Dichev & Dicheva, 2017; Mekler et al., 2017). Hence, the goal of this study was to compare the effectiveness of badges and leaderboards both individually and together on student academic performance and motivation by implementing them in two online classes. Two experiments were conducted, and the grading systems of two classes were either partially or fully gamified by implemented badges and leaderboards.

2 Literature review

2.1 Theoretical background

The goal-setting theory by Locke and Latham (2002) is commonly referred to by gamification scholars to understand how gamification tools affect the performance and motivation of users (Dichev & Dicheva, 2017; Tondello et al., 2018).

Goals can be defined as end-states that a person wants to achieve within a certain timeframe in specific contexts (Zimmerman & Schunk, 2008). Setting a goal is suggested to influence people's motivation and performance through four mechanisms: providing cognitive and behavioral directions, increasing people's energies and efforts to handle tasks, increasing persistence to complete tasks, and evoking affective reactions, such as increased satisfaction (Locke & Latham, 2006; Zimmerman, 2008). Hence, gamification tools might help people set goals (especially when rules are clearly identified) and increase their goal-related behaviors for achievement (Fanfarelli et al., 2015; Hakulinen et al., 2013; Hamari, 2017; Morris et al., 2019; Tondello et al., 2018). Through these four mechanisms explained above, gamification tools could encourage users to set goals to pursue, to direct their attention and effort to accomplish the task, to increase their persistence with the try-fail-try again feature, and to evoke positive affection, such as the feeling of competence and self-efficacy after goal accomplishment (Tondello et al., 2018). The goals in the gamified environment could be both explicitly set, such as going for a quest, or implicitly set such as earning badges or ranking at a higher spot in a leaderboard as an outcome of an activity (Morris et al., 2019; Tondello et al., 2018). Landers et al. (2015) gamified a classic brainstorming task with a leaderboard and assigned participants randomly into five groups. Four of these groups were classic goal-setting levels: do-your-best, easy, difficult, and impossible goals. The fifth group was a leaderboard group, where the participants were not instructed to achieve any goal. The mere presence of a leaderboard motivated participants to perform at similar levels to participants in difficult and impossible goal-setting groups, which was interpreted as participants setting goals to be at or near the top of the leaderboard.

Feedback is another mechanism for gamification tools to affect the performance and motivation of users (Dichev & Dicheva, 2017; Hamari, 2017; Kapp, 2012; McDaniel & Fanfarelli, 2015). Feedback is defined as information about how a person's current state of knowledge and performance relates to set goals and standards (Hattie & Timperley, 2007), and it has an essential role in the performance and motivation of learners (Burgers et al., 2015; Fyfe & Rittle-Johnson, 2016; Shin & Dickson, 2010). Gamification tools may deliver performance and mastery feedback to students, which in turn may affect their future performance and motivation, such as by evoking feelings of competence (Mekler et al., 2017; Sailer et al., 2017). It was found that the feedback provided by gamification tools is recognized and appreciated by students (Alabbasi, 2017; Cheong et al., 2014).

2.2 The effects of gamification on performance and motivation

There are mixed findings regarding the effects of gamification on performance and motivation in educational environments (Buckley & Doyle, 2016; Dichev & Dicheva, 2017; Seaborn & Fels, 2015). For example, in one experiment students in the gamified group were exposed to badges, leaderboards, and trophies while completing course activities, and they scored higher on practical assignments and reported positive attitudes toward the use of gamification tools. However, those

same students performed worse on written assignments and participated less compared to the control group (de-Marcos et al., 2014; Domínguez et al., 2013). Frost et al. (2015) implemented badges, points, leaderboard, narration, and lives in a learning management system. They investigated the effects of those gamification tools on outcomes of interest, motivation, satisfaction, learning (measured by grades), and perception of pedagogical affect. They found that none of the outcomes were affected significantly by gamification tools, except the motivation and interest (although with small effect sizes). Despite these nonsignificant effects, students reported that they liked the gamification aspects of the course (Frost et al., 2015). On the contrary, Hanus and Fox (2015) gamified a lesson with implemented badges, a leaderboard, and coins and found that students in the gamified course showed less intrinsic motivation, satisfaction, empowerment, and lower final exam scores than those in the non-gamified class. In addition to these mixed and negative results, some studies found strong positive effects of gamification tools on student performance and motivation. For example, gamification tools have been implemented within Massive Open Online Courses (MOOCs) to increase student engagement (Chang & Wei, 2016). In one of these implementations, the inclusion of gamification tools along with social media increased participation, learning motivation, and learning the course content while increasing the completion rate from an average rate of 7–39.9% (Borras-Gene et al., 2016).

When the COVID-19 pandemic disrupted traditional education and students lived under social isolation, gamification was considered as a prominent candidate method to overcome students' lack of motivation and engagement issues (Nair, 2021; Rincon-Flores & Santos-Guevara, 2021; Zainuddin et al., 2021). The different gamification approaches that were used during the pandemic have been reviewed by Nieto-Escamez and Roldán-Tapia (2021). These studies generally showed that most students approached gamification positively as an innovative, engaging, effective method for online courses. Despite the positive findings on the motivation and engagement of students, none of the reviewed studies found an objective improvement in students learning due to gamification. In addition, a few studies included in the review did not find a positive effect of gamification on students' motivation and performance, possibly due to the negative moods of students during confinement. Other studies, which were not included in the review, also reached a similar conclusion as Nieto-Escamez and Roldán-Tapia (2021). To illustrate, da Silva Junior et al. (2022) gamified two online undergraduate classes using points, badges, leaderboards, rewards, and educational games. According to survey results, students approached gamification with high positive attitudes; however, the effect of gamification on students' performance was inconclusive. The enhanced engagement and motivation effects of gamification on students during the pandemic were also supported by other studies (Al Breiki & Yahaya, 2021; Chans & Portuguese Castro, 2021; Nair, 2021; Rincon-Flores & Santos-Guevara, 2021; Zainuddin et al., 2021). A few studies also found improved student academic performance (grades) due to gamification (Chans & Portuguese Castro, 2021; Rincon-Flores & Santos-Guevara, 2021). Hence, conflicting results about the effects of gamification on motivation and performance before and during the pandemic necessitate the need for further research (Dichev & Dicheva, 2017; Nieto-Escamez & Roldán-Tapia, 2021).

Badges and leaderboards are the most frequently implemented gamification tools (Hamari et al., 2014; Seaborn & Fels, 2015). Despite their popularity, there is no strong evidence for the effectiveness of badges and leaderboards (Dichev & Dicheva, 2017). Hence, badges and leaderboards were chosen in this study to compare their effectiveness on academic performance and motivation. An overview of research results about badges and leaderboards will be further discussed in more detail (see Table 1 for the summary of the studies on badges and leaderboards).

2.3 Badges

Badges are digital rewards given for accomplishments, which also carry information about users' mastery and performance levels (Abramovich et al., 2013). In addition, badges can function as immediate feedback given to users about their performance after completing a task (Kapp, 2012; McDaniel & Fanfarelli, 2015). Furthermore, when clearly identified rules of earning badges are given, they can help students set goals for themselves and encourage goal-related behaviors (Hamari, 2017; Sailer et al., 2017). Badges are used as a graphical icon in gamification studies, and they usually become visible after users accomplish specific tasks as an acknowledgment of their skills and achievements (da Rocha Seixas et al., 2016; Grant, 2014). Users usually follow their earned badges on their personalized badge page (Denny, 2013).

Studies that implemented badges in online environments yielded somewhat positive but still mixed results (see Table 1). Badges have been proposed to have the potential to decrease well-known student problems in online education, such as procrastination and motivation problems (Haaranen et al., 2014; Hakulinen & Auvinen, 2014). Hakulinen et al. (2015) found that badges affected student behaviors positively in terms of more time spent per exercise, more number of sessions, and more total time spent in an online class. Also, their survey revealed positive attitudes toward badges by most students. McDaniel et al. (2012) used badges as one part of the grading system to encourage students for early submission before deadlines and for providing helpful feedback to peers. Student attitudes for the badge system were moderately positive; however, many students reported frustration due to hidden and hard-to-find badges, which decreased their favorability. In another study, Denny (2013) implemented badges into the PeerWise platform, in which students created and answered questions. In this study, badges motivated students in the treatment group to contribute more answers (but not more questions) and spend more time with the system. The survey given after the course showed that most students preferred the system with badges and liked to see badges at their user interfaces. Kyewski and Krämer (2018) implemented badges in an online class in which badges were awarded to students based on quiz scores and some class activities, such as participating in the discussion board or providing peer feedback. However, they found that badges did not affect students' motivation, course engagement, and academic performance (quiz scores and final grades), regardless of whether badges were only visible to students themselves or to both students and peers. Finally, two experiments implemented badges, learning goals, and badges with learning goals in low- and high-stakes learning contexts and found no effects of these elements on

Table 1 Studies on badges and leaderboards

Study	Gamification tool	Purpose of gamification	Outcome
Denny (2013)	Badges	Encouraging participation to classwork	Partially increased contributions and more extended engagement with course work. Students prefer the system with badges
Hakulinen et al. (2015)	Badges	Encouraging engagement with classwork	Increased motivation and desired study practices
Kyewski and Krämer (2018)	Badges	Encouraging engagement with classwork Enhancing motivation and performance	No improvement on motivation, engagement, or performance of students
McDaniel et al. (2012)	Badges and Leaderboards	Encouraging targeted student behaviors	Students liked the gamified class design, but their preference for badges was modestly positive
Morris et al. (2019)	Badges	Encouraging learning	No improvement in learning
Banfield and Wilkerson (2014)	Leaderboards	Enhancing student motivation and self-efficacy toward course content	Increased motivation and self-efficacy
Hanus and Fox (2015)	Badges and Leaderboards	Enhancing desired student behaviors, satisfaction, empowerment, and academic performance	No improvement of motivation, satisfaction, empowerment, or academic performance
Landers and Landers (2014)	Leaderboards	Encouraging targeted student behaviors	Increased time on task and student interaction
Landers et al. (2015)	Leaderboards	Improving task performance	Improved task performance

learning outcomes (Morris et al., 2019). In conclusion, students tend to evaluate the use of badges positively; however, the effect of badges on student performance was not clear due to the mixed findings.

2.4 Leaderboards

Leaderboards are digital rankings of students based on their performances. Leaderboards provide individual-level feedback by reporting personal accomplishment and progress, while they provide group-level feedback by enabling comparisons with the performance of others (Landers & Landers, 2014; Nebel et al., 2017). Similar to badges, leaderboards may encourage students to set goals for themselves and increase their performance (Landers et al., 2015). For example, Landers and Landers (2014) found that the addition of leaderboards into course design significantly increased the interaction between students and their projects compared to students who did not see leaderboards.

Some researchers criticized leaderboards due to the possible adverse effect of leaderboard-prompted social comparisons on student behavior and motivation (Hanus & Fox, 2015, please refer to Table 1). However, others proposed that competition caused by leaderboards can have a constructive effect on participation and learning through social comparison (Banfield & Wilkerson, 2014; Sailer et al., 2017). There is currently no consensus on whether the leaderboard has a positive (opportunity to track progress among other students) or negative effects (harmful effect of competition) on motivation and performance (Hung, 2017).

The social comparison caused by leaderboards could be understood through the social comparison theory of Festinger (1954). This theory states that people continually compare themselves with others as this is a fundamental psychological mechanism that affects people's judgments, experiences, and behavior. People engage in social comparisons, especially in times of uncertainty and novelty, because they need to maintain a stable and accurate self-view from informative feedback they receive about their characteristics and abilities through comparisons (Festinger, 1954; Wan & Sadiq, 2012). People benefit from objective standards or choose other people similar to them as comparison standards to gain an accurate self-evaluation (Corcoran et al., 2011; Michinov & Primois, 2005). Hence, digital leaderboards could be possible tools for online students to self-evaluate by comparing their performance with similar others, their peers in the online class. In addition to enabling self-evaluation, social comparison also serves other functions based on people's current motivations (Wan & Sadiq, 2012). For instance, people engage in social comparison with worse-off others (downward comparison) with the concern of self-enhancement, or they may compare themselves with better-off others (upward comparison) with the concern of self-improvement (Garcia et al., 2013). That said, people are more prone to compare themselves with others who are slightly better than them (upward comparison) (Christy & Fox, 2014; Festinger, 1954) to motivate themselves to improve their performance with higher set goals (Mechi, & Sanchez-Mazas, 2012; Michinov & Primois, 2005). For example, people increased their productivity and creativity in online group activity when they had the opportunity to compare their contributions

to those of other group members (Michinov & Primois, 2005). Hence, leaderboards could be a tool for online students to engage in upward or downward comparison based on their current motivations.

2.5 Current study

Instructors in higher education favor gamification to increase student attention and learning, enable interactive learning, and motivate students through entertainment (Sanchez-Mena & Marti-Parreno, 2017). However, despite instructors' expectations, gamification does not always lead to enhanced performance and motivation, as noted above. The conflicting results of gamification studies make it necessary for further research to understand the conditions under which gamification enhances motivation and performance, which is the first aim of this paper (Dicheva et al., 2015; Faiella & Ricciardi, 2015; Hung, 2017).

One reason for inconsistent results in the gamification literature is that gamified environments can be created in countless ways by implementing various combinations of gamification tools (Dichev & Dicheva, 2017; Seaborn & Fels, 2015). More importantly, most gamification studies implemented more than one gamification tool simultaneously, limiting the utility of the results (Dichev & Dicheva, 2017). That is, the relative and additive contribution of each gamification tool remains unclear without a clear experimental design (Dichev & Dicheva, 2017). Hence, two gamification tools, badges and leaderboards, were chosen for this purpose as they are the most frequently used gamification tools. We investigated the effects of badges and leaderboards separately and in combination on academic performance and motivation, which is the second aim of this study (Hamari et al., 2014; Mekler et al., 2017; Seaborn & Fels, 2015). To the best of our knowledge, this study is among the first studies which compare the relative effects of badges and leaderboards on academic performance and motivation (Hamari et al., 2014; Mekler et al., 2017).

The effects of badges and leaderboards were investigated on two outcomes: academic performance and motivation. As shown in Table 1, the effects of badges and leaderboards on academic performance and motivation are contradictory, which necessitate further research. Academic performance of students was defined as quiz scores of students in Experiment 1 and both quiz and assignment scores (which was also the final course grade) in Experiment 2. The second outcome, motivation, was defined as motivational beliefs and attitudes of students towards badges and leaderboards (Haaranen et al., 2014; Fotaris et al., 2016). Thus, the first research question we investigated was: *How do badges and leaderboards differ in terms of their effects on academic performance and self-reported motivation levels of online students when they are implemented individually versus together?*

The third aim of the study is to investigate the effect of leaderboard-prompted comparisons on academic performance and motivation, as the few empirical studies conducted on the topic resulted in conflicting results (Hanus & Fox, 2015; Landers & Landers, 2014). Further research is needed due to the lack of consensus on whether leaderboards positively (e.g., an opportunity for self-evaluation) or negatively (e.g., the harmful effect of competition) affect the motivation and performance

of online students (Hung, 2017). Thus, our second research question was: *Does leaderboard-prompted social comparison affect student performance and motivation positively or negatively?* Applying Festinger's social comparison theory (1954), we also investigated if students engaged in upward or downward comparison when exposed to digital leaderboards in an online learning environment (Corcoran et al., 2011).

The fourth aim of the study is to address the methodological problems that gamification studies have been criticized for, such as lack of control groups or brief durations (Hamari et al., 2014; Seaborn & Fels, 2015). This paper addresses these issues by gamifying semester-long two online courses. Also, we used a true experimental design by randomly assigning participants into three experimental groups and one control group.

The present study also aims to contribute to the existing literature by gamifying a given course to varying degrees, partially- gamified grading system (Experiment 1) and fully-gamified grading system (Experiment 2) (Hung, 2017). Hence, we compared differences in performance and motivational outcomes caused by different gamification designs by conducting two experiments, which is the fifth aim of this study. Each of the five aims of this study addresses the gaps in the gamification literature in which further research is needed.

3 Experiment 1

Experiment 1 assessed whether badges and/or leaderboards were effective in increasing student academic performance and motivation in a partially-gamified online class. Gamification tools were implemented in only the quiz section of the grading system, which made up 40% of the course grade.

3.1 Method

3.1.1 Participants

One hundred nine students gave their consent for this experiment; however, seven students dropped the course. Thus, 102 undergraduate students were recruited from an online introductory undergraduate physics course in a Northeastern Ohio public university during the Spring 2016 semester. The study was approved by the Institutional Review Board (IRB) of the university. The subject information survey was completed by 68 participants: there were 56 (82%) female participants, 67 (99%) participants were in the age range of 18–25, and all participants were over 18 years of age. All participants were given 2 extra credits (0.85% of total course grade) to participate in this experiment. Convenience sampling was used for this study, which is a non-probability type of sampling where participants are included in the study as they are easy to reach by the researchers (Wiersma & Jurs, 2009).

Leaderboards	Badges	
	Yes	No
Yes	Badges with Leaderboard Group	Leaderboard-Only Group
No	Badges-Only Group	No Badges No Leaderboard (Control) Group

Fig. 1 2 x 2 factorial design



Fig. 2 Phases of the study

3.1.2 Materials and procedure

The undergraduate physics course used in Experiment 1 was fully online and was taught through Blackboard Learn System. All registered students were invited to the study via an email sent by the instructor. The experimental design was a 2 (badges vs. no badges) x 2 (leaderboards vs. no leaderboards) factorial design (please see Fig. 1). Students who agreed to participate were randomly assigned to one cell of a 2x2 design: (a) badges-only group, (b) leaderboards-only group, (c) badges with leaderboards group, and (d) no badges and no leaderboards (control) group. Four different groups were created on the course's Blackboard site. The course content and grading system were the same for all groups. Figure 2 shows the phases of the study.

Implemented badges into Blackboard were accessible for badges-only and badges with leaderboards groups (will be referred to as “badges groups”), and uploaded leaderboards into Blackboard were available for leaderboards-only and badges with leaderboards groups (will be referred to as “leaderboards groups”). The control group was not exposed to badges or leaderboards, and they were only asked to submit the subject information survey and the motivation survey.

Students in the badges groups (badges-only and badges with leaderboards groups) were awarded badges based on their quiz performance through the “Achievements” course tool of Blackboard. Badge images were designed by the investigators (see Fig. 3). The course had 11 quizzes prepared by the course instructor, and each student could earn a badge in each of the quizzes. Eight quizzes had 10 multiple choice questions, while the other three had five multiple choice questions. There were three



Fig. 3 Images for gold, silver and bronze badges

levels for each badge: bronze, silver, and gold. Three difficulty levels were chosen for badges to increase students' self-efficacy and encourage them to earn higher level badges (Banfield & Wilkerson, 2014). In addition, three difficulty levels of the badge system will foster students' persistence by having them set higher level goals for themselves (Tondello et al., 2018) and provide them a sense of progression towards mastery (de-Marcos et al., 2014). In quizzes with 10 questions, students scoring 5–6, 7–8, and 9–10 received a bronze, silver, and gold badge, respectively. In quizzes with five questions, scores of 3, 4, and 5 represented bronze, silver, and gold, respectively. A link to a page named “Your Badges!” was created on the course's home page for only the badges groups. If participants earned a badge after finishing a quiz, they were notified with a green banner at the top of the page to inform them about the badge they earned. Participants could see all earned badges on their “Your Badges!” page.

As for the leaderboard groups (leaderboards-only and badges with leaderboards groups), students were ranked based on their total quiz scores. Due to the lack of a built-in leaderboard application in Blackboard, leaderboards were created manually and posted as a “.pdf” file in Blackboard after the quiz deadline. The posted leaderboards were visible only to leaderboard groups. Leaderboards were created based on the total score students earned on all quizzes to that point in the semester. For example, the third leaderboard was created based on participants' aggregated scores from the first, second, and third quizzes. Students in the leaderboard groups were assigned a pseudonym from among Nobel Prize-winning physicists' last names, and participants were informed of their pseudonym by email. Only these pseudonyms were used in the leaderboards. Different leaderboard formats were used for leaderboards-only and badges with leaderboards groups. That is, leaderboards for the leaderboards-only group consisted of pseudonyms, ranks, and aggregated total scores. As for badges with leaderboards group, in addition to pseudonyms, ranks, and aggregated total score, earned badge images for each quiz were also displayed on the leaderboards, as shown in Fig. 4.

The rules of how to earn gold, silver, and bronze badges were explained to students in badges groups. Similarly, students in leaderboard groups were explained how the leaderboards were created. The rules were explained to students with the
















Pseudonyms	Rank	Quiz 1-3	Quiz 1	Quiz 2	Quiz 3	Pseudonyms	Rank	Quiz 1-3
Davis	1	30				Kajita	1	30
Koshiba	1	30				McDonald	1	30
Giacconi	1	30				Amano	1	30
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Perl	7	19		None		Hall	10	8
Taylor	8	16			None	Politzer	11	7
Richardson	9	15			None	Wineland	12	0

Fig. 4 Examples of Posted Leaderboards for Badges with Leaderboards Group (Left) and Leaderboards-only Group (Right)

purpose of encouraging them to set goals and increase their goal-related behaviors for earning badges and ranking higher in the leaderboards (Hamari, 2017; Landers et al., 2015; Sailer et al., 2017). Besides, badges and leaderboards could offer extrinsic motivational support for goal achievement when students earn a badge or are ranked high at the top of a leaderboard (Hakulinen & Auvinen, 2014; Mekler et al., 2017; Sailer et al., 2017).

In this online class, students were allowed to take all quizzes multiple times via Blackboard and the highest score from all attempts was recorded as the final quiz score by Blackboard. The quiz questions were shuffled for each quiz attempt. This allowed us to examine the performances of students in three domains for each quiz: highest score received from all attempts, first attempt score, and the number of attempts. As eight quizzes had 10 multiple choice questions and three quizzes had five multiple choice questions, we doubled the score obtained from these three quizzes to bring all the quiz scores to the same scale.

Two surveys were collected from participants. First, the subject information survey was collected at the beginning of the study and the participants were asked their age, gender, and major. Besides their demographic information, two questions were asked to learn the gaming experience of participants. We asked the participants if they like playing video games or social network games. If their answer was “yes” to playing video games, they were also asked the approximate number of hours they play games in a day. The gaming experience could be an important confounding variable, which may affect the attitudes and motivation of participants towards the badges and leaderboards (Cheong et al., 2014; Hanus & Fox, 2015). Second, students’ motivation and attitudes towards badges and leaderboards were measured by a self-report survey. This motivation survey was collected through Blackboard at the end of the study. Survey questions for badges were based on the questions used by Haaranen et al. (2014) and modified for the current study (please see Table 3 for survey questions of badges groups). The authors created all survey questions for leaderboard groups (Table 4 shows the survey questions for leaderboard groups). A five-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) was used for

the motivation survey. One open-ended question was also included in the survey to receive additional comments from participants about badges and/or leaderboards. Participants in the badges groups responded to questions about badges and those in the leaderboards groups responded to leaderboard-related questions. Participants in badges with leaderboards group answered both badge-related questions and leaderboard-related questions. Each of the surveys has a good internal consistency: Cronbach's alpha was 0.92 for the badges motivation survey and was 0.86 for the leaderboard motivation survey. Participants in the control group were asked different but related questions, such as if they would have preferred having badges and/or leaderboards implemented in the course design.

3.2 Results and discussion

3.2.1 Performance on quizzes

We investigated the effects of badges and leaderboards on quiz performances and motivation of online students. A 2×2 repeated-measures multivariate analysis of variance (MANOVA) was conducted to examine gamification tools' differences in quiz performance in three domains: highest score received from all attempts, first attempt scores, and the number of attempts. No significant differences in the quiz outcomes were found for any of the between-subject effects: badges (Wilk's Lambda=0.984, $F(3, 96)=0.535$, $p=.660$), leaderboards (Wilk's Lambda=0.982, $F(3, 96)=0.597$, $p=.618$), or the interaction of badges and leaderboards (Wilk's Lambda=0.963, $F(3, 96)=1.24$ $p=.30$). However, the within-subject variable (i.e. 11 quizzes) was significant: Wilk's Lambda=0.047, $F(30, 69)=46.709$ $p=.00$. This significant difference between quizzes was not of interest for the purposes of the study. The averages for the highest scores, the initial scores, and the number of attempts from 11 quizzes were calculated. These mean scores and standard deviations for three dependent variables for each group were included in Table 2. The MANOVA results indicated that gamification tools did not affect the obtained highest scores, the initial scores, and the number of attempts for 11 quizzes throughout the semester. No matter which group they were in, the students were mostly trying to reach the highest score with multiple attempts, which caused a ceiling effect in the

Table 2 Average of descriptive statistics in quiz performance for Experiment 1

Groups	<i>n</i>	Highest Score	Initial Score	Number of Attempts
		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
BO	26	9.43 (1.2)	4.57 (1.2)	4.15 (1.5)
LO	26	9.23 (0.96)	4.37 (1.4)	4.02 (1.5)
BWL	25	9.52 (0.54)	4.95 (1.4)	3.94 (1.1)
Control	25	9.51 (0.45)	4.66 (1.2)	4.47 (1.7)
Total	102	9.42 (0.82)	4.64 (1.3)	4.15 (1.4)

Note. BO=Badges-only group, LO=Leaderboard-only group, BWL=Badges with Leaderboard group. The maximum score for any quiz was 10.

quiz performance and resulted in similar results for all groups. Such nonsignificant variations among groups were in line with some studies (Frost et al., 2015; Kyewski & Krämer, 2018).

It was found that being exposed to two gamification tools (both badges and leaderboards) or only one gamification tool (either badges or leaderboards) did not result in significant differences in the three domains of quiz performances. Moreover, there was no difference in the number of earned gold, silver, and bronze badges between badges-only and badges with leaderboards groups. Due to a Blackboard-related technical problem, the log data for the number of views of participants to the “Your Badges!” page and posted leaderboards were lost for this experiment. Hence, the manipulation check to control if participants followed their badges and/or looked at leaderboards could not be conducted.

3.2.2 Subject information survey

In the subject information survey ($n=68$), in addition to their age and gender, which were reported in the Participants section, participants were also asked if they liked playing video games or social network games. Forty-six participants (67.6%) answered “yes” to this question, while 22 participants (32.4%) answered “no”. When asked about the approximate number of hours they play such games in a day, 25% of students replied less than 1 h, 14.7% one hour, 17.6% 2 h, 5.9% 3 to 4 h, 4.4% 5+ hours, and while 32.4% did not play.

The Chi-square test showed no group differences for both questions, i.e., whether they like playing games or how many hours they play games ($p>.05$). However, a gender difference was found for how many hours participants play games, $\chi^2(10)=22.27$, $p=.014$. A greater proportion of males than expected played games 1 h daily compared to females. There was no gender difference for other categories of the number of hours spent on games daily.

3.2.3 The motivation survey

Participants were asked about their attitudes and motivation toward the gamification tools at the end of the semester. Sixty-eight participants (badges-only group=19, leaderboards-only group=16, badges with leaderboards group=15 and control=18) responded to this survey. The means, standard deviations, and percentages of answers for survey questions are shown in Table 3 for badges groups and Table 4 for leaderboard groups. In the survey, the participants were also asked one open-ended question about badges and/or leaderboards, the analysis of which is provided in Appendix 1.

The composite motivation scores were computed by taking the average of items in the badges motivation survey. The mean motivation score was 3.68 ($SD=1.04$) for the badges-only group and 3.78 ($SD=0.89$) for badges with leaderboards group. No significant difference between groups was found in the t-test for independent groups on their responses to the badges motivation survey, $t(32)=-0.294$, $p=.77$. As for the leaderboard motivation survey, the mean motivation score was 3.52 ($SD=0.80$) for the leaderboards-only group and 3.45 ($SD=0.76$) for badges with

Table 3 Motivation survey questions for badges in Experiment 1

Questions	Groups	<i>M (SD)</i>	1	2	3	4	5
I like the implemented badges in this class	BO	4 (.91)	0%	0%	42.1%	21.1%	36.8%
	BWL	3.7 (.82)	0%	6.7%	33.3%	46.7%	13.3%
The badges were motivating for me	BO	3.5 (1.3)	10.5%	5.3%	36.8%	15.8%	31.6%
	BWL	3.8 (1.2)	0%	13.3%	33.3%	13.3%	40%
Availability of badges had a positive impact on my behavior	BO	3.4 (1.3)	10.5%	10.5%	36.8%	15.8%	26.3%
	BWL	3.7 (1.1)	0%	13.3%	33.3%	20%	33.3%
I think that badges should be used in other online classes as well	BO	3.9 (.99)	0%	0%	52.6%	5.3%	42.1%
	BWL	3.9 (.87)	0%	0%	40%	26.7%	33.3%

Note. BO = Badges-only group, BWL = Badges with leaderboard group. 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree

Table 4 Motivation survey questions for leaderboards in Experiment 1

Questions	Groups	<i>M (SD)</i>	1	2	3	4	5
I like the implemented leaderboards in this class	LO	3.9 (1.0)	0%	6.3%	37.5%	18.8%	37.5%
	BWL	3.7 (.96)	0%	6.7%	40%	26.7%	26.7%
Having the leaderboards in this class was motivating for me	LO	3.8 (1.3)	6.3%	12.5%	12.5%	31.3%	37.5%
	BWL	3.5 (1.1)	0%	20%	33.3%	26.7%	20%
Knowing my rank among other students made me work harder	LO	3.6 (1.3)	6.3%	12.5%	31.3%	18.8%	31.3%
	BWL	3.6 (.98)	0%	13.3%	33.3%	33.3%	20%
Leaderboard facilitated monitoring my progress relative to other students	LO	3.6 (1.2)	6.3%	6.3%	37.5%	25%	25%
	BWL	3.6 (.93)	0%	6.7%	40%	26.7%	20%
Comparing my rank with other students was discouraging for me	LO	2.8 (1.3)	18.8%	25%	25%	18.8%	12.5%
	BWL	2.4 (.83)	6.7%	60%	20%	13.3%	0%
I compared myself with other students who ranked higher than me in the leaderboard	LO	3.1 (1.3)	12.5%	18.8%	25%	31.3%	12.5%
	BWL	3.1 (1.0)	6.7%	20%	40%	26.7%	6.7%
I compared myself with other students who ranked lower than me in the leaderboard	LO	2.9 (1.2)	12.5%	25%	31.3%	18.8%	12.5%
	BWL	2.9 (.92)	6.7%	26.7%	40%	26.7%	0%
I think that leaderboards should be used in other online classes as well	LO	4.1 (1.2)	0%	6.3%	18.8%	37.5%	37.5%
	BWL	3.6 (1.2)	6.7%	6.7%	33.3%	26.7%	26.7%

Note. LO = Leaderboards-only group, BWL = Badges with leaderboard group. 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree

the leaderboards group. The t-test for independent groups showed no group differences for motivation scores of leaderboards groups, $t(29)=0.233$, $p=.82$.

A one-way ANOVA showed no difference between males and females in their motivation score for the badges and leaderboards surveys ($p>.05$). The badges and leaderboard motivation scores were not significantly different for participants who like playing games, who do not like playing games, and who play games for a varying number of hours ($p>.05$).

As shown in Tables 3 and 4, the mean scores of positively worded questions in the motivation survey mainly were between 3 (Neutral) and 4 (Agree). Hence, on average, the participants showed positive attitudes toward gamification tools, agreeing with results in the gamification literature (de-Marcos et al., 2014; Denny, 2013; Domínguez et al., 2013; Hakulinen et al., 2015). Furthermore, either implemented alone or together, badges and leaderboards yielded similar positive ratings in the motivation survey. Contrary to Hanus and Fox (2015), results from both the quiz performances and motivation survey show that the badges and leaderboards, implemented alone or together, did not negatively affect students' quiz performances and motivation throughout the semester. In addition, most students reported that they liked the badges and leaderboards and wanted to see them in other online classes.

It has been claimed that leaderboards could lead to competition, negatively affecting users (Christy & Fox, 2014; Hanus & Fox, 2015). To test this concern, participants were asked to rate the question of "Comparing my rank with other students was discouraging for me". As shown in Table 4, 43.8% of participants in the leaderboards-only group and 67.7% of participants in badges with leaderboards group responded either "Strongly Disagree" or "Disagree" to this question, while a smaller percentage of participants (31.3% for leaderboards-only and 13.3% for badges with leaderboards) responded "Strongly Agree" or "Agree". Hence, the majority of the participants in badges with leaderboards group (67.7% vs. 13.3%) were not affected by the negative impact of the comparison through the leaderboards. However, close percentage rates (43.8% vs. 31.3%) for the leaderboards-only group showed that leaderboards' negative effect was more apparent for this group. The number of gamification tools the students were exposed to could explain this difference between the badges with leaderboards group and the leaderboards-only group. Implementing only leaderboards could be discouraging for students as they are the only point of comparison and feedback form for students, while implementing leaderboards with badges would be less problematic as they also earn badges, which are another form of feedback about their performance.

According to social comparison theory, people compare themselves to different people for different purposes. They engage in downward comparisons for self-enhancement and upward comparisons for self-improvement (Garcia et al., 2013). To test what type of comparisons participants made, they were asked if they compared their rank to those who ranked lower or higher than themselves. As shown in Table 4, 43.8% of participants (those who chose "Agree" and "Strongly Agree") in leaderboards-only reported that they compared themselves to those who ranked higher (upward comparison), while 31.3% stated they compared themselves to those who ranked lower than their rank (downward comparison). As for badges with leaderboards group, 33.4% engaged in an upward comparison, while 26.7% made

a downward comparison. As the upward comparison percentages were higher than those for downward comparisons for both groups, the participants seemed to engage in upward comparisons more frequently. This finding is similar to literature in that people are more prone to make an upward comparison to motivate themselves with higher set goals (Christy & Fox, 2014; Festinger, 1954; Mechi & Sanchez-Mazas, 2012; Michinov & Primois, 2005). In addition, approximately half of the students in both leaderboard groups reported using leaderboards to monitor their progress, and knowing their ranks encouraged them to work harder.

The control group was also asked about their motivation level throughout the semester and if they would prefer having badges and leaderboards implemented in the course design. Eighteen control group participants answered the motivation survey. When they were asked if they liked the design of the course, 88.9% of participants responded “Agree” and “Strongly Agree” to this question. When they were asked if they were motivated throughout the course, 77.8% responded “Agree” and “Strongly Agree”. The students were also asked if they prefer having a leaderboard in the course: slightly more students agreed with this proposition (Strongly Disagree: 16.7%, Disagree: 16.7%, Neutral: 27.8%, Agree: 33.3%, Strongly Agree: 5.6%). They were also asked if they prefer having the implemented badge system in the course, and higher favorable ratings compared to leaderboard were received for this question: (Strongly Disagree: 5.6%, Disagree: 5.6%, Neutral: 44.4%, Agree: 33.3%, Strongly Agree: 5.6%).

In conclusion, the results yielded that gamification tools were not effective on performance, but they positively affected online students’ motivation. One possible reason for the nonsignificant effects of badges and leaderboards in quiz performance could be that only 40% of the grading system was gamified, resulting in an incomplete representation of users’ mastery and performance feedback (Abramovich et al., 2013). The nonsignificant result obtained in this study might be due to the inability of gamification tools to encourage students to set higher goals for themselves as they were implemented in only 40% of the grading system. Hence, gamifying the whole grading system could enhance the effectiveness of the gamification tools to set higher goals and provide feedback about progress. To research this question, the second experiment was conducted in which all course evaluation was gamified. Another reason for the nonsignificant effect could be that the online class that was gamified might not be the best option to observe the effects of gamification tools on the performance and motivation of students. The control group’s motivation survey revealed that most students liked the current course design and were motivated throughout the semester. Hence, another introductory undergraduate physics online class taught by the same instructor was used for Experiment 2.

4 Experiment 2

Experiment 2 assessed the effects of badges and/or leaderboards in a fully-gamified online class. Experiment 2 used the same design as Experiment 1, with the exceptions noted in the Materials and Procedure section below. The modifications in the experiment were approved by the university IRB.

4.1 Method

4.1.1 Participants

One hundred and one undergraduate students from an online introductory physics course in a Northeastern Ohio public university gave their consents to participate in Experiment 2. However, 10 students dropped the course and were excluded from the study. Two participants were excluded from the analysis because they completed less than one-fifth of the course requirements. One student was excluded from analysis, as this student was younger than 18 years old. Thus, 88 participants formed the final subject sample: badges-only ($n=20$), leaderboards-only ($n=22$), badges with leaderboards ($n=23$), control ($n=23$). The subject information survey was completed through Blackboard by 83 participants. The mean age of participants was 20.67 ($SD=4.84$), and all participants were over 18 years of age. There were 53 females (63.9%), 28 males (33.7%), and 2 (2.4%) participants preferred not to answer. All participants were given 10 extra credits (3.7% of total course grade) for their participation in this experiment.

4.1.2 Materials and procedure

The procedure used in Experiment 1 was also used in Experiment 2 (please see Fig. 2). Hence, participants were randomly assigned into four cells of 2×2 factorial design (badges-only group, leaderboards-only group, badges with leaderboards group, and no badges and no leaderboards [control] group) (please refer to Fig. 1). The subject information survey was collected from all groups at the beginning of the semester. Participants were exposed to badges if they were in badges-only and badges with leaderboard groups throughout the semester. And participants in the leaderboards-only group and badges with leaderboards group were exposed to the leaderboards throughout the semester. The participants in the control group were exposed to neither badges nor leaderboards. Finally, the motivation survey was collected at the end of the semester as the last step of the study. Convenience sampling was used; hence participants were recruited from the students' population that the researchers have access to (Wiersma & Jurs, 2009).

Despite the overall similarities to Experiment 1, eight changes were made for Experiment 2. (1) A different fully online introductory undergraduate course offered in the physics department during the Spring 2017 semester was used. This course was taught by the same instructor and had a similar difficulty level to the course used in Experiment 1. (2) The whole course grade, which included 9 assignments (66.66% of course grade) and 9 quizzes (33.33% of course grade), was gamified. Badges were given based on both quiz and assignment scores, and leaderboards were created based on the total cumulative scores of quizzes and assignments up to that point. (3) There were five modules in the course, and a new leaderboard was uploaded after each module, instead of after each quiz as it was done in Experiment 1. So, a total of five leaderboards were posted in Experiment 2, while 11 were posted in Experiment 1. (4) Reminder emails about the posted gamification tools were sent to participants after completion of each module by the instructor and the

investigators. (5) Participants in the leaderboard groups were allowed to choose their pseudonyms. If they did not choose a pseudonym, investigators assigned a pseudonym from among the last names of Nobel Prize-winning physicists. (6) One type of leaderboard format was used for both leaderboard groups, which consisted of pseudonyms, ranks, and aggregated total scores. (7) The badge earning mechanism was the same as the first experiment. Students were allowed to take quizzes multiple times, and the highest score from all their attempts for a particular quiz was their final score for that quiz. The students could earn one badge for each quiz, so badges were replaced when they earned a higher-level one. However, students were allowed to submit their assignments only once for grading. Students earned a gold badge if the assignment score was 19 or 20 out of 20, a silver badge if the assignment score was 17 or 18, or a bronze badge if the assignment score was 15 or 16. (8) As stated above, there were five modules in this course. The earned quiz and assignment scores in each module were added to create a module score. These five module scores and final course grades were used as indicators of academic performance in the second experiment, while only quiz performance was used in the first experiment. The same self-report motivation survey was collected to measure students' motivation and attitudes towards badges and leaderboards. Cronbach's alpha showed that the badges motivation survey used in Experiment 2 has a good internal consistency, $\alpha=0.90$. The leaderboard motivation survey also has an adequate internal consistency: Cronbach's $\alpha=0.78$.

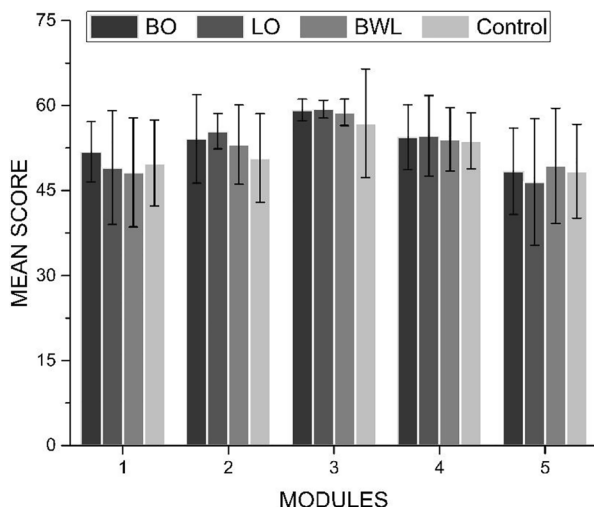
4.2 Results and discussion

4.2.1 Performance on module scores and final course grade

A two-way mixed ANOVA was conducted to compare the module scores of groups across five modules to examine if gamification tools impacted participants' academic performance. The results indicated that there were significant differences across five modules as a within-subjects factor, Wilk's $\Lambda=0.43$, $F(4, 81)=26.68$, $p<.001$, $\eta^2=0.57$ (Fig. 5), however this significant difference between modules was not of interest for this study. The badges main effect was nonsignificant, Wilk's $\Lambda=0.99$, $F(4, 81)=0.22$, $p=.93$, as well as the leaderboards main effect, Wilk's $\Lambda=0.94$, $F(4, 81)=1.24$, $p=.30$. Finally, the interaction of badges and leaderboards was also nonsignificant, Wilk's $\Lambda=0.95$, $F(4, 81)=1.14$, $p=.35$. Although the control group scored lowest in Modules 2, 3, and 4 (Fig. 5), these differences were nonsignificant.

A two-way ANOVA was conducted to examine group differences for the final course scores of participants. There was no significant difference for the badges, $F(1, 84)=0.37$, $p=.55$, nor for the leaderboards, $F(1, 84)=0.00$, $p=.99$. The interaction term between the badges and the leaderboards was also nonsignificant, $F(1, 84)=0.77$, $p=.38$. Besides, there was no difference in the number of earned gold, silver, and bronze badges between badges-only and badges with leaderboards groups.

Fig. 5 Mean scores of groups across five modules. Note: BO=Badges-only group, LO=Leaderboard-only group, BWL=Badges with leaderboard group. The maximum score was 60 for each module



The analyses showed that overall course performances of students in the badges groups and leaderboard groups did not differ significantly from the performances of control group students, similar to Experiment 1. Also, we replicated the findings of the first experiment that the badges and leaderboards, either implemented alone or together, did not differ in terms of effectiveness and they did not have any negative effect on the performances of students throughout the semester, despite some claims (Hanus & Fox, 2015).

4.2.2 Number of views for gamification tools

Statistical tracking features of Blackboard enabled us to track the number of times the participants viewed the badge page (in which they monitored their earned badges) and the posted leaderboards. This analysis was not performed for Experiment 1, as these data were lost due to a Blackboard-related technical problem. As explained in the procedure section of Experiment 1, students can follow their earned badges on the “Your Badges!” page that can be accessed from the home page of the online class. Number of views for the badges was compared for badges-only group ($M=44.5$, $SD=18.82$) and badges with leaderboards group ($M=40.96$, $SD=11.18$). A one-way ANOVA showed no significant difference between groups, $F(1, 41)=0.58$, $p=.45$. Although participants could earn a maximum of 18 badges from both quizzes and assignments, both groups looked at the badge page more than necessary, as mean numbers for both groups were above 40. Students’ motivation and interest could explain this high number of views for badges, even though earning badges did not affect their performances.

Similarly, the number of views for the posted five leaderboards was tracked for leaderboards-only and badges with leaderboards groups. The mean number of views for five posted leaderboards were as follows: leaderboards-only group $M=10.41$, $SD=7.15$ and badges with leaderboards group $M=7.7$, $SD=8.54$. A one-way

ANOVA showed no significant difference among leaderboards-only and badges with leaderboards groups in the number of views for posted five leaderboards, $F(1, 43) = 1.33$, $p = .26$. In addition, 6 students in badges with leaderboards group and 3 students in leaderboards-only group never looked at any posted leaderboards, which caused a caveat for the fidelity issue for these groups. The manipulation check was conducted by removing the students who never looked at the leaderboards from the subjects' pool. We then re-ran the analyses for performance and motivation variables after removing these students. The overall results for academic performance and motivation did not change: the two-way mixed ANOVA, which was conducted to compare groups in the module scores across five modules, showed that both badges main effect, Wilk's Lambda = 0.99, $F(4, 73) = 0.20$, $p = .94$, and leaderboards main effect, Wilk's Lambda = 0.96, $F(4, 73) = 0.77$, $p = .55$, were nonsignificant. Finally, the interaction of badges and leaderboards was also nonsignificant, Wilk's Lambda = 0.96, $F(4, 73) = 0.74$, $p = .57$. In addition, the two-way ANOVA, which was conducted to examine group differences for the final score, showed no significant difference for the badges, $F(1, 76) = 0.00$, $p = .99$, nor for the leaderboards, $F(1, 76) = 0.63$, $p = .43$. The interaction term between the badges and the leaderboards was nonsignificant as well, $F(1, 84) = 2.51$, $p = .12$. The overall results for the leaderboard motivation survey were similar to the results when all participants were included in the data.

Both leaderboard groups looked at five posted leaderboards more than necessary, similar to badges groups, as the mean number of views for both groups was more than 5 times. A Spearman's rank-order correlation was run to determine the relationship between the number of views for leaderboards and the ranks of the students in the leaderboards. There was no significant correlation for badges with leaderboards group for all five leaderboards. However, there were two significant correlations for the leaderboards-only group. That is, there was a moderate, negative correlation between students' ranks in leaderboard 1 and their number of views for leaderboard 1, $r_s = -0.52$, $p = .013$. Also, there was a moderate, negative correlation between students' ranks in leaderboard 2 and their number of views for leaderboard 2, $r_s = -0.48$, $p = .023$. There were no significant correlations for the third, fourth, and fifth leaderboards in the leaderboards-only group. Hence, students who ranked higher in the first and second leaderboards in the leaderboards-only group looked at the posted leaderboards more often than those who ranked lower.

4.2.3 The subject information survey

Eighty-three participants filled out this survey, in which they were asked their age and gender, which were reported in the Participants section. In addition, they were asked if they like playing video games or social network games. 55 participants (66.3%) answered "yes" to this question, while 28 participants (33.7%) answered "no". When asked how many hours approximately they play such games in a day, 33.7% of them played less than one hour, 13.3% played 1 h, 8.4% played 2 h, 10.8% played 3 to 4 h, 3.6% played more than 5 h, and 30.1% did not play. A Chi-square test showed no group differences for both questions (i.e., whether they like playing games ($p = .13$) or how many hours they play games ($p = .10$)). However, gender

differences were found for whether they like playing games, $\chi^2(2) = 11.94, p = .003$. Fewer males than expected by chance reported that they did not like playing games. Another gender difference was found for the time spent daily playing games, $\chi^2(10) = 26.55, p = .003$. Fewer females than expected by chance played games 3–4 h daily, and more males than expected by chance played games 3–4 h daily. These results suggest that more males reported liking playing games and playing for longer hours than females.

4.2.4 The motivation survey

The same motivation survey used in Experiment 1 was collected at the end of the semester, and it was filled out by 81 participants (badges-only = 18, leaderboards-only = 20, badges with leaderboards = 22, Control = 21). The means, standard deviations, and percentages of answers for survey questions are shown in Table 5 for badges groups and Table 6 for leaderboards groups. One open-ended question was asked about badges and leaderboards, the analysis of which is reported in Appendix 2.

The composite motivation scores were computed by taking the average of items in the badges motivation survey. The mean motivation score was 3.58 ($SD = 0.83$) for badges-only group and it was 3.84 ($SD = 0.83$) for badges with leaderboards group. The t-test for independent groups was run for the badges motivation survey to examine group differences for motivation scores of badges groups. There was no significant difference between the groups on their responses to the badges motivation survey, $t(38) = -0.98, p = .33$. As for the leaderboard motivation survey, the mean motivation score was 3.53 ($SD = 0.62$) for the leaderboards-only group, and it was 3.80 ($SD = 0.57$) for the badges with leaderboards group. The t-test for independent groups showed no group differences for motivation scores of leaderboards groups, $t(40) = -1.51, p = .14$.

No gender difference was found for badges and leaderboards motivation scores based on a one-way ANOVA ($p > .05$). The badges and leaderboards motivation scores were not significantly different for participants who like to play games, who do not like to play games, and who play games on a varying number of hours ($p > .05$).

The survey questions for the control (no badges, no leaderboard) were different from the questions of experimental groups. When they were asked if they liked the regular course design (in which no gamification tools were implemented), all of them responded “Agree” and “Strongly Agree” to this question. When they were asked if they were motivated throughout the course, 76.2% of the participants responded either “Agree” or “Strongly Agree”. Thus, the current course design without any added gamification tools was rated positively and evaluated as motivating for students, similar to Experiment 1. Students were also asked whether they would have preferred having a leaderboard in the course. Most students responded negatively to this suggestion (Strongly Disagree: 19%, Disagree: 33.3%, Neutral: 19%, Agree: 19%, Strongly Agree: 9.5%). They were also asked if they would have preferred having the implemented badge system in the course and the distribution

Table 5 Motivation survey questions for badges in Experiment 2

Questions	Groups	<i>M (SD)</i>	1	2	3	4	5
I like the implemented badges in this class	BO	3.9 (.68)	0%	0%	27.8%	55.6%	16.7%
	BWL	4 (.82)	0%	0%	31.8%	36.4%	31.8%
The badges were motivating for me	BO	3.2 (1.2)	5.6%	22.2%	38.9%	16.7%	16.7%
	BWL	3.6 (1.1)	0%	18.2%	31.8%	22.7%	27.3%
Availability of badges had a positive impact on my behavior	BO	3.4 (1.0)	0%	22.2%	27.8%	33.3%	16.7%
	BWL	3.8 (.97)	0%	9.1%	31.8%	31.8%	27.3%
I think that badges should be used in other online classes as well	BO	3.8 (.79)	0%	5.6%	22.2%	55.6%	16.7%
	BWL	4 (.87)	0%	4.5%	22.7%	40.9%	31.8%

Note. BO = Badges-only group, BWL = Badges with leaderboard group. 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree

Table 6 Motivation survey questions for leaderboards in Experiment 2

Questions	Groups	<i>M (SD)</i>	1	2	3	4	5
I like the implemented leaderboards in this class	LO	3.8 (.77)	0%	5%	25%	55%	15%
	BWL	4.1 (.72)	0%	0%	22.7%	45.5%	27.3%
Having the leaderboards in this class was motivating for me	LO	3.6 (.94)	0%	15%	25%	45%	15%
	BWL	4 (.93)	0%	9.1%	13.6%	45.5%	31.8%
Knowing my rank among other students made me work harder	LO	3.5 (1.1)	5%	15%	15%	55%	10%
	BWL	4 (.89)	0%	4.5%	27.3%	36.4%	31.8%
Leaderboard facilitated monitoring my progress relative to other students	LO	3.9 (.64)	0%	0%	25%	60%	15%
	BWL	4 (.98)	0%	9.1%	18.2%	36.4%	36.4%
Comparing my rank with other students was discouraging for me	LO	2.5 (1.1)	15%	45%	30%	0%	10%
	BWL	2.9 (1.1)	0%	50%	27.3%	9.1%	13.6%
I compared myself with other students who ranked higher than me in the leaderboard	LO	3.4 (.99)	0%	20%	40%	25%	15%
	BWL	3.7 (.99)	0%	13.6%	22.7%	40.9%	22.7%
I compared myself with other students who ranked lower than me in the leaderboard	LO	3 (1.1)	5%	30%	40%	10%	15%
	BWL	3.6 (1.1)	0%	18.2%	27.3%	36.4%	18.2%
I think that leaderboards should be used in other online classes as well	LO	3.6 (1.2)	10%	5%	25%	40%	20%
	BWL	4.1 (.79)	0%	4.5%	13.6%	54.5%	27.3%

Note. LO = Leaderboards-only group, BWL = Badges with leaderboard group. 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree

of responses was as follows: Strongly Disagree: 9.5%, Disagree: 23.8%, Neutral: 23.8%, Agree: 28.6%, Strongly Agree: 14.3%.

The motivation survey responses were similar to the first experiment: the participants' attitudes toward badges and leaderboards were on average positive, as the mean scores of positively worded questions were between 3 (Neutral) and 4 (Agree). Students' number of views based on log data for badges and leaderboards also supported their positive attitudes toward gamification tools. Participants liked both tools, found them motivating and encouraging to work harder. Gamification tools were also desired to be included in future online classes.

As in Experiment 1, we also analyzed the positive and negative effects of social comparison caused by leaderboards. As shown in Table 6, more than 60% of the students in both leaderboard groups reported using leaderboards to monitor their progress and knowing their ranks encouraged them to work harder. When asked if comparing their rank with other students was discouraging for them, 10% of participants in the leaderboards-only group and 22.7% of participants in the badges with leaderboards group responded "Agree" or "Strongly Agree", while the majority were not affected. In addition, it was found that 40% of participants in the leaderboards-only group (those who chose "Agree" or "Strongly Agree") reported that they compared themselves to those who ranked higher (upward comparison), while 25% stated they compared themselves to those who ranked lower than themselves (downward comparison). As for badges with leaderboards group, 63.6% engaged in an upward comparison, while 54.6% made a downward comparison. As the percentages for upward comparison were higher for both groups, the participants seemed to engage in upward comparisons more frequently, similar to Experiment 1. This result is in line with the literature in that people are more likely to make an upward comparison to motivate themselves with higher set goals (Festinger, 1954; Mechi & Sanchez-Mazas, 2012; Michinov & Primois, 2005).

5 General discussion

This empirical study investigated several research questions by conducting two experiments. The first research question was whether badges and leaderboards differed in effectiveness on academic performance and motivation of students in online classes. To the best of our knowledge, this study is among the first to investigate the relative effects of two frequently used gamification tools, badges and leaderboards, on academic performance and motivation (Hamari et al., 2014; Mekler et al., 2017). This study also contributed to the existing gamification literature by gamifying an online class in varying degrees (Hung, 2017). The grading system of an online class was partially gamified in the first experiment and fully gamified in the second experiment; which enabled us to compare the effects of different gamification designs.

In both experiments, there was no significant difference between badges and leaderboards in terms of their effectiveness on the academic performances of online students compared to the control group. In addition, there was no additional benefit of including both tools in one course as implementing two gamification tools did not improve student performance. Moreover, the two experiments demonstrated

that gamifying an online course partially or fully did not result in different levels of effectiveness, and no significant improvement in student academic performance was observed in either condition. Therefore, our results were similar to other studies, in which no significant positive or negative effect of gamification tools was found on student behaviors (Frost et al., 2015; Kyewski & Krämer, 2018; Morris et al., 2019).

The results are somewhat surprising because gamification tools can theoretically enhance many aspects of performance. For example, badges and leaderboards could provide performance and mastery feedback (McDaniel & Fanfarelli, 2015; Mekler et al., 2017) and set clear performance goals for students (Hamari, 2017; Landers et al., 2015). In our experiments, even though the rules of earning gold badges and ranking higher in the leaderboards were explained to students, these gamification tools did not result in any additional performance gains for the experimental groups compared to the control group for both experiments. This could be due to the inability of the current implementation style of gamification tools to provide performance feedback or to enable students to set clear performance goals for themselves. As there was no guidance from the instructor to students about how to benefit from gamification tools as a performance feedback mechanism for setting higher goals, the students might have ignored these advantages of gamification tools.

Although gamification tools were not effective on academic performance, they were associated with positive attitudes in most of the students in both experiments, consistent with previous research (de-Marcos et al., 2014; Denny, 2013; Domínguez et al., 2013; Hakulinen et al., 2015). The implementation of badges and leaderboards, either individually or together, resulted in similar ratings in terms of motivation and positive attitudes in both experiments. Most participants reported that they liked the badges and/or leaderboards, found them encouraging and motivating, and preferred to see them in future online classes. Similar to our findings, studies conducted during the COVID-19 pandemic found strong positive effects of gamification on online students' motivation (Chans & Portuguese Castro, 2021; da Silva Junior et al., 2022; Rincon-Flores & Santos-Guevara, 2021). Hence, these researchers suggested that gamification could be used as an innovative and effective method to support online students' motivation in the post-COVID-19 era as well (Nieto-Escamez & Roldán-Tapia, 2021; Rincon-Flores & Santos-Guevara, 2021).

The second research question of this study was to understand whether leaderboard-prompted social comparison affects student performance and motivation positively or negatively in online classes. Both experiments concluded that leaderboards were rated positively. That is, more students agreed than disagreed with the statements of the motivation survey that leaderboards facilitated monitoring their progress relative to others, and knowing their ranks encouraged them to work harder. These favorable ratings could be interpreted as the leaderboard's ability to provide clear goals to students and motivate them extrinsically to be at higher ranks in the leaderboards, in addition to providing feedback about their status among others (Landers et al., 2015). Also, when asked about the possible harmful effects of leaderboards due to social comparison and competition, the majority of the participants reported that they were not negatively affected by these comparisons (contrary to Hanus & Fox, 2015). The negative effects of leaderboards were probably minimized, as only pseudonyms were used in the leaderboards, and due to the online nature of classes, students did not encounter one another

in person. However, a smaller percentage of students reported being discouraged by comparison and competition caused by the leaderboards. To avoid such instances, being excluded from the leaderboards could be an option for students. According to Festinger's (1954) social comparison theory, people engage in upward comparisons for self-improvement and downward comparisons for self-enhancement (Corcoran et al., 2011). In this context, both experiments yielded similar results in that participants engaged in upward comparisons more frequently than downward comparisons, which is in line with the literature (Christy & Fox, 2014).

There could be several possible explanations for the obtained nonsignificant results for academic performance. First, the two online classes gamified for this study might not be the ideal classes for the study's aims. That is, the majority of control group participants in both experiments reported in the motivation survey that they liked how the online classes were set up and they were motivated throughout the semester. In addition, the high success rates for these two introductory online classes were also apparent with the students' final grades: 77% of the participants in Experiment 1 and 69% of the participants in Experiment 2 obtained the final grade of A or A- from the courses. Apparently, the highly experienced instructor in addition to the well-organized course designs left little room to improve student performances by gamification tools in these courses. Second, badges and leaderboards are not effective tools to improve the academic performance of online students. However, we approach this explanation cautiously, as it contradicts previous positive findings. A more cautious approach would be to state that the types of badges and leaderboards did not enhance student performance under the conditions and implementation style of our experiments (see Morris et al., 2019 for similar findings). Third, we only added gamification tools in the grading systems of two courses without gamifying the course content. This type of gamification could have limited the potential impact of gamification tools on student behaviors, as gamification tools might not result in a game-like experience for students or add an additional challenge for students to pursue their self-set goals (Hung, 2017; Seaborn & Fels, 2015). Besides, earning badges and being in the top ranks of the leaderboards did not affect the grading of the students, which may make some students indifferent to them, as stated in their answers for the open-ended question in the motivation survey. Additionally, gamification tools might not provide adequate extrinsic motivational support to increase performance, as they did not affect grading (Hakulinen & Auvinen, 2014; Mekler et al., 2017). In conclusion, the nonsignificant results of this study could be due to one or more reasons cited above. An online class could be gamified in various ways, and our way of implementation of gamification tools in the online classes chosen for this study might have fallen short of creating the conditions to reveal their effectiveness (Dichev & Dicheva, 2017; Sailer et al., 2017).

6 Practical Implications

The results of this study provide valuable practical implications to the researchers and the instructors who want to gamify their courses. First, researchers should check the previous class evaluations and feedback from students before gamifying

an online class to see if gamification would resolve any of the issues raised by students or if there is a need for gamification. These evaluations might also lead to what type of gamification tool would serve best to improve the course. Our two experiments showed that if the students are already satisfied with the non-gamified online class, then gamification tools may have a limited effect on academic performance. Moreover, we gamified only the grading system (partially or fully) without presenting any additional challenge to students to follow (Haaranen et al., 2014; da Rocha Seixas et al., 2016). Limiting the gamification only to performance assessment may not be enough to benefit from gamification tools' feedback and goal setting features. Hence, additional challenges connected to the gamification tools based on the course content might be created (Mekler et al., 2017; Morris et al., 2019). In our study, students did not earn any score from earning badges or ranking higher in the leaderboards, so the gamification tools only had a symbolic value. Although some researchers warned that the gamification system should not be tied to the course grading system (Haaranen et al., 2014), they may at least provide some bonus points to the students (Kyewski & Krämer, 2018). Tying bonus points to the number of earned badges or ranking high in the leaderboards may provide an additional challenge to students and increase the value of the gamification tools in their eyes.

7 Limitations and Future Research

There are several limitations for this study. First, the two chosen online classes were introductory undergraduate physics courses which were among the core courses required by the university. Hence, future work should test if these results are consistent for elective courses, courses from different departments, and different levels, including graduate-level online classes (Seaborn & Fels, 2015). Second, the effects of gamification tools were investigated with undergraduate students in a U.S. university. Competitive environment caused by leaderboard may be more tolerable for U.S. students (Lander & Landers, 2014), while different results may be reached in other cultures which empathize collaboration within the classroom instead of competition. Hence, future research should explore if there are potential cultural differences in the effectiveness of gamification tools, especially the leaderboards. Third, we focused on two common gamification tools, badges and leaderboards. Hence, our findings are limited to these two gamification tools, and nonsignificant results in academic performance may not be valid for other gamification tools. Fourth, we did not provide virtual social environments to participants to share and show their earned badges, which could be one method to increase participants' interest in gamification tools (McDaniel et al., 2012). Fifth, the badges and leaderboards were designed and created by the authors in this study, which may decrease the overall effectiveness of both gamification tools as they might be perceived as simple and dull designs with a lack of visual charm and appeal. Hence, future researchers might benefit from a professional graphic designer to design more attractive badges and leaderboard images. Lastly, as the primary purpose was to measure students' motivation and attitudes towards badges and leaderboards, we only used a self-report survey with questions dedicated to two gamification tools. However, a validated measurement

for motivation could have been included to measure the overall motivation of all students toward the course, which may enable comparison between the experimental and control groups.

A suggestion for future studies will be that researchers should check the previous class evaluations before gamifying an online class to see if the course needs improvements with gamification tools. Some leaderboard formats could be more appealing for students than the format used in this study, where students saw a static leaderboard for each quiz or module. One possible design could be a live leaderboard that will update the rank of students immediately after they earn a score. The live leaderboard design may provide immediate feedback to students about their progress, which would be an essential improvement for the leaderboard design since the timeliness of feedback was rated more important than the extent of feedback by online students (Dennen et al., 2007). Another possible design for the leaderboard tool could be using “team leaderboards”, in which teams will be ranked based on their task performance. Hence, not only cooperation within a group but also competition among groups may motivate online students. Researching the effectiveness of different formats for the leaderboard tool would be a possible next step in this field.

8 Conclusions

In this study, several contributions to the gamification literature have been made. First, our results demonstrated that when badges and leaderboards were implemented in the grading system of the online classes, there was no significant improvement or deterioration in the academic performance of students compared to the control group. This nonsignificant result did not change when gamification tools were implemented partially or fully in the grading system. When they were implemented individually or together, badges and leaderboards yielded similar results in terms of effectiveness on academic performance and motivation. We also found that leaderboard-prompted social comparison did not affect student performance negatively, and leaderboards oriented them to make the upward comparison more frequently. Finally, most students perceived badges and leaderboards positively, and students expressed their desire to see them in future online classes.

Appendix 1

Participants were asked one open-ended question to share their additional comments about badges and/or leaderboards at the end of the motivation survey in Experiment 1. Fourteen participants provided comments about badges ($n=9$ badges-only group and $n=5$ for badges with leaderboards group). These responses were categorized into five themes: (a) liked and motivated by badges, (b) liked but not necessarily motivated by badges, (c) promising if better implemented, (d) indifferent, and (e) others. Three participants (approx. 21%) reported clearly in their answers that they liked and were motivated by the badges. For example, one participant wrote: “*I liked the badges personally, because it motivated me to get the best score I could possibly*

get”. While some students ($n=5$, approx. 35%) reported that they liked badges, they also included comments about not being motivated much by badges or they did not comment about motivation at all. For example, one student wrote: *“The badges were nice but I wasn’t that motivated to work harder. I worked hard to get 100% on the quizzes and didn’t think about the badges”*. Two participants (approx. 14%) commented that they were not interested in badges in this class but said they could be a promising tool if implemented differently. One student wrote: *“I didn’t think they were that important because I didn’t get notified about getting one. I think to make the game version more fun is to get notified about receiving a badge”*. Furthermore, three participants (approx. 21%) wrote that they were indifferent toward badges: *“I honestly didn’t notice them that much”*. Lastly, one answer (approx. 7%), which did not fit any of the categories, was grouped into an “others” category. This answer was: *“The badges give an idea to the student of how well they are doing throughout the course of the semester”*.

The same categorization system was also used for the leaderboards. Thirteen participants ($n=6$ from badges with leaderboards group and $n=7$ from the leaderboards-only group) provided feedback about the leaderboards. Of those, three participants (approx. 23%) stated that they liked and were motivated by the leaderboards. For example, one participant responded: *“I found the leaderboard to be a great reference tool and it was very motivating!”*. One participant (approx. 7%) declared that she/he liked leaderboards but the response (*“I forgot about it at times, but I did like the leaderboard!”*) did not mention motivation. Two participants (approx. 15%) stated that leaderboards would be promising after improvements in implementation. For example, one of these participants wrote: *“To be honest, I really didn’t check it that often because I couldn’t figure out which name I was. But I feel like if I did know it would of been kinda cool to see if I was doing better than others”*. Six participants (approx. 46%) were indifferent toward leaderboards. An example response from one of these participants was: *“I never looked at the leaderboard”*. Finally, the following answer (approx. 7%) was categorized in “others” category: *“I will have to retake this course due to not being able to put in the adequate enough time to finish it”*.

Appendix 2

Participants were asked to share their additional comments about badges and leaderboards with an open-ended question at the end of the motivation survey in Experiment 2. Twenty-nine participants provided comments for badges ($n=15$ from the badges-only group and $n=14$ from badges with leaderboards group). These responses were categorized into the following five themes (similar to Experiment 1): (a) liked and motivated by badges, (b) liked but not necessarily motivated by badges, (c) promising if better implemented, (d) indifferent, and (e) others. Eight participants (approx. 27%) reported clearly in their answers that they liked and were motivated by the badges. For example, one participant wrote: *“I really enjoyed them. Even though it was small, it was motivating to earn them and pushed me to try again for a higher grade”*. While some students ($n=9$, approx. 31%) reported that they liked badges, they also included not being motivated much by badges or they

did not comment about motivation at all. For example, one student wrote: *“The badges didn’t motivate me to do my work, but I liked the concept of the badges”*. Four participants (approx. 13%) commented that although they were not interested in badges in this class, badges could be a promising tool if they would be implemented differently. One student wrote: *“I almost forgot they were there. It could be helpful and motivational to students more if they are more aware of them. The way it is set up currently doesn’t really have the desired impact”*. Furthermore, six participants (approx. 20%) wrote that they were indifferent toward badges. For example, one student wrote: *“I didn’t really pay much attention to them”*. Lastly, two answers (approx. 6%) did not fit in any of the five categories and were grouped in a category called “others”. One of these comments was the following: *“The Badges didn’t motivate me as much as the leader board did”*.

Twenty-seven participants ($n=11$ from badges with leaderboards group and $n=16$ from leaderboards-only group) provided feedback about the leaderboards. The same categorization system was used for the leaderboards. Of those, thirteen participants (approx. 48%) stated that they both liked and were motivated by the leaderboards. An example response for this category was the following: *“I liked having the leaderboard. It allowed me to track how I was doing relative to others and motivated me to be on top”*. Eight participants (approx. 29%) declared that they liked leaderboards, but it was not necessarily motivating for them or they did not mention motivation in their responses. For example, one student responded as follows: *“I think that the leaderboard was fun however, I do not think it really encouraged me to work harder than I would have without it”*. Three participants (approx. 11%) stated that leaderboards would be promising after improvements in implementation. An example response for this category was the following: *“... I would rather have a constantly updating leader board instead of one for each act [module], I’d like to see one with live updates...”*. Two participants (approx. 7%) were indifferent toward leaderboards and one of them responded as follows: *“I actually forgot it was an option. I wasn’t concerned where I ranked in the class and merely did my best”*. Finally, the following response (approx. 4%) was categorized in “others” category: *“I get the appeal of competitiveness in school, but it doesn’t seem to blend as well into the class than something like sports”*

Authors’ contributions Sebiha Balci: Conceptualization, Methodology, Formal analysis, Investigation, Writing - Original Draft, Review & Editing, Project administration.

Jonathan M. Secaur: Data collection, Project administration.

Bradley J. Morris: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Project administration.

All authors read and approved the final manuscript.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publish All participants gave their consents regarding publishing their data.

Research involving human participants and/or animals Approval was obtained from the ethics committee of Kent State University. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

References

- Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education?: It depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, 61(2), 217–232. <https://doi.org/10.1007/s11423-013-9289-2>
- Al Breiki, M., & Yahaya, W. A. J. W. (2021). Using gamification to promote students' engagement while teaching online during COVID-19. In I. Fayed & J. Cummings (Eds.), *Teaching in the post COVID-19 era* (pp. 443–453). Springer. https://doi.org/10.1007/978-3-030-74088-7_44
- Alabbasi, D. (2017). Exploring graduate students' perspectives towards using gamification techniques in online learning. *Turkish Online Journal of Distance Education*, 18(3), 180–196
- Alaswad, Z., & Nadolny, L. (2015). Designing for game-based learning: The effective integration of technology to support learning. *Journal of Educational Technology Systems*, 43(4), 389–402. <https://doi.org/10.1177/0047239515588164>
- Banfield, J., & Wilkerson, B. (2014). Increasing student intrinsic motivation and self-efficacy through gamification pedagogy. *Contemporary Issues in Education Research*, 7(4), 291–298
- Borras-Gene, O., Martinez-Nunez, M., & Fidalgo-Blanco, A. (2016). New challenges for the motivation and learning in engineering education using gamification in MOOC. *International Journal of Engineering Education*, 32(1), 501–512
- Buckley, P., & Doyle, E. (2016). Gamification and student motivation. *Interactive Learning Environments*, 24(6), 1162–1175. <https://doi.org/10.1080/10494820.2014.964263>
- Burgers, C., Eden, A., Van Engelenburg, M. D., & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior*, 48, 94–103. <https://doi.org/10.1016/j.chb.2015.01.038>
- Chang, J. W., & Wei, H. Y. (2016). Exploring engaging gamification mechanics in massive online open courses. *Educational Technology & Society*, 19(2), 177–203. <https://www.jstor.org/stable/jeductechsci.19.2.177>
- Chans, G. M., & Portuguese Castro, M. (2021). Gamification as a strategy to increase motivation and engagement in higher education chemistry students. *Computers*, 10(10), Article 132. <http://doi.org/10.3390/computers10100132>
- Cheong, C., Filippou, J., & Cheong, F. (2014). Towards the gamification of learning: Investigating student perceptions of game elements. *Journal of Information Systems Education*, 25(3), 233–245
- Christy, K. R., & Fox, J. (2014). Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women's math performance. *Computers and Education*, 78, 66–77. <https://doi.org/10.1016/j.compedu.2014.05.005>
- Corcoran, K., Crusius, J., & Mussweiler, T. (2011). Social comparison: Motives, standards, and mechanisms. In D. Chadee (Ed.), *Theories in social psychology* (pp. 119–139). Wiley-Blackwell
- Da Rocha Seixas, L., Gomes, A. S., & De Melo Filho, I. J. (2016). Effectiveness of gamification in the engagement of students. *Computers in Human Behavior*, 58, 48–63. <https://doi.org/10.1016/j.chb.2015.11.021>
- da Silva Júnior, J. N., de Lima Castro, G., Melo Leite Junior, A. J., Monteiro, A. J., & Alexandre, F. S. O. (2022). Gamification of an entire introductory organic chemistry course: A strategy to enhance the students' engagement. *Journal of Chemical Education*. Advance online publication. <https://doi.org/10.1021/acs.jchemed.1c00766>

- De-Marcos, L., Domínguez, A., Saenz-De-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers and Education*, 75, 82–91. <https://doi.org/10.1016/j.compedu.2014.01.012>
- Dennen, V., & Bonk, C. J. (2007). We'll leave the light on for you: Keeping learners motivated in online courses. In B. H. Khan (Ed.), *Flexible learning in an information society* (pp. 64–76). The Idea Group, Inc
- Dennen, V. P., Darabi, A. A., & Smith, L. J. (2007). Instructor-learner interaction in online courses: The relative perceived importance of particular instructor actions on performance and satisfaction. *Distance Education*, 28(1), 65–79. <https://doi.org/10.1080/01587910701305319>
- Denny, P. (2013). The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI conference on human factors in computing systems, New York, USA* (pp. 763–772). <https://doi.org/10.1145/2470654.2470763>
- Dennon, A. (2021). *Coronavirus impacts on students and online learning*. BestColleges. Retrieved December 3, 2021 from <https://www.bestcolleges.com/blog/coronavirus-impacts-on-students/>. Accessed 12 Feb 2021.
- Deterding, S., Khaled, R., Nacke, L., & Dixon, D. (2011). Gamification: Toward a definition. In *Proceedings of the CHI 2011 gamification workshop, Vancouver, BC, Canada*
- Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education*, 14(9), 1–36. <https://doi.org/10.1186/s41239-017-0042-5>
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Educational Technology & Society*, 18(3), 75–88. <https://doi.org/10.1109/EDUCON.2014.6826129>
- Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers and Education*, 63, 380–392. <https://doi.org/10.1016/j.compedu.2012.12.020>
- Faiella, F., & Ricciardi, M. (2015). Gamification and learning: A review of issues and research. *Journal of e-Learning and Knowledge Society*, 11(3), 13–21
- Fanfarelli, J., Vie, S., & McDaniel, R. (2015). Understanding digital badges through feedback, reward, and narrative: A multidisciplinary approach to building better badges in social environments. *Communication Design Quarterly Review*, 3(3), 56–60. <https://doi.org/10.1145/2792989.2792998>
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117–140. <https://doi.org/10.1177/001872675400700202>
- Fotaris, P., Mastoras, T., Leinfellner, R., & Rosunally, Y. (2016). Climbing up the leaderboard: An empirical study of applying gamification techniques to a computer programming class. *Electronic Journal of E-Learning*, 14(2), 94–110. <https://doi.org/10.1007/s11528-016-0070-3>
- Frost, R. D., Matta, V., & MacIvor, E. (2015). Assessing the efficacy of incorporating game dynamics in a learning management system. *Journal of Information Systems Education*, 26(1), 59–70
- Fyfe, E. R., & Rittle-Johnson, B. (2016). The benefits of computer-generated feedback for mathematics problem solving. *Journal of Experimental Child Psychology*, 147, 140–151. <https://doi.org/10.1016/j.jecp.2016.03.009>
- Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science*, 8(6), 634–650. <https://doi.org/10.1177/1745691613504114>
- Ginder, S.A., Kelly-Reid, J.E., and Mann, F.B. (2019). Enrollment and employees in postsecondary institutions, fall 2017; and financial statistics and academic libraries, fiscal year 2017: First look (provisional data) (NCES 2019-021rev). U.S. Department of Education. Washington, DC: National Center for Education Statistics. <https://nces.ed.gov/pubs2019/2019021REV.pdf>. Accessed 16 Jan 2021.
- Grant, S. (2014). Badges: Show what you know. *Young Adult Library Services*, 12(2), 28–32. <https://doi.org/10.1177/1098214014565505>
- Haaranen, L., Ihantola, P., Hakulinen, L., & Korhonen, A. (2014). How (not) to introduce badges to online exercises. In *Proceedings of the 45th ACM technical symposium on computer science education, Atlanta, USA* (pp. 33–38). <https://doi.org/10.1145/2538862.2538921>
- Hakulinen, L., & Auvinen, T. (2014). The effect of gamification on students with different achievement goal orientations. In *Proceedings of the 2014 international conference on teaching and learning in computing and engineering, Kuching, Malaysia* (pp. 9–16). <https://doi.org/10.1109/LaTiCE.2014.10>

- Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). Empirical study on the effect of achievement badges in TRAKLA2 online learning environment. In *Proceedings of the learning and teaching in computing and engineering (LaTiCE), Macau* (pp. 47–54). <https://doi.org/10.1109/LaTiCE.2013.34>
- Hakulinen, L., Auvinen, T., & Korhonen, A. (2015). The effect of achievement badges on students' behavior: An empirical study in a university-level computer science course. *International Journal of Emerging Technologies in Learning*, 10(1), 18–30. <https://doi.org/10.3991/ijet.v10i1.4221>
- Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, 71, 469–478. <https://doi.org/10.1016/j.chb.2015.03.036>
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii international conference on system sciences, Hawaii, USA* (pp. 3025–3034). <https://doi.org/10.1109/HICSS.2014.377>
- Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers and Education*, 80, 152–161. <https://doi.org/10.1016/j.compedu.2014.08.019>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hung, A. C. (2017). A critique and defense of gamification. *Journal of Interactive Online Learning*, 15(1), 57–73
- Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Pfeiffer
- Kywewski, E., & Krämer, N. C. (2018). To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. *Computers & Education*, 118, 25–37. <https://doi.org/10.1016/j.compedu.2017.11.006>
- Landers, R. N., Bauer, K. N., & Callan, R. C. (2015). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior*, 71, 508–515. <https://doi.org/10.1016/j.chb.2015.08.008>
- Landers, R. N., & Landers, A. K. (2014). An empirical test of the theory of gamified instructional design: The effect of leaderboards on academic performance. *Simulation & Gaming*, 45, 769–785. <https://doi.org/10.1177/1046878114563662>
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57, 705–717. <https://doi.org/10.1037/0003-066X.57.9.705>
- Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, 15(5), 265–268. <https://doi.org/10.1111/j.1467-8721.2006.00449.x>
- Mahle, M. (2011). Effects of interaction on student achievement and motivation in distance education. *Quarterly Review of Distance Education*, 12(3), 207–215
- McDaniel, R., & Fanfarelli, J. R. (2015). A digital badging dataset focused on performance, engagement and behavior-related variables from observations in web-based university courses. *British Journal of Educational Technology*, 46(5), 937–941. <https://doi.org/10.1111/bjet.12272>
- McDaniel, R., Lindgren, R., & Friskics, J. (2012). Using badges for shaping interactions in online learning environments. In *Proceedings of the 2012 IEEE international professional communication conference, Orlando, USA* (pp. 1–4). <https://doi.org/10.1109/ipcc.2012.6408619>
- Means, B., Bakia, M., & Murphy, R. (2014). *Learning online: What research tells us about whether, when and how*. Routledge Taylor & Frances
- Mechi, A., & Sanchez-Mazas, M. (2012). Who strives and who gives up? The role of social comparison distance and achievement goals on students' learning investment. *Problems of Education in the 21st Century*, 48, 107–116
- Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, 71, 525–534. <https://doi.org/10.1016/j.chb.2015.08.048>
- Michinov, N., & Primois, C. (2005). Improving productivity and creativity in online groups through social comparison process: New evidence for asynchronous electronic brainstorming. *Computers in Human Behavior*, 21(1), 11–28. <https://doi.org/10.1016/j.chb.2004.02.004>
- Morris, B. J., Dragovich, C., Todaro, R., Balci, S., & Dalton, E. (2019). Comparing badges and learning goals in low- and high-stakes learning contexts. *Journal of Computing in Higher Education*, 31(3), 573–603. <https://doi.org/10.1007/s12528-019-09228-9>

- Nair, B. B. (2021). Endorsing gamification pedagogy as a helpful strategy to offset the COVID- 19 induced disruptions in tourism education. *Journal of Hospitality, Leisure, Sport & Tourism Education*. Advance online publication. <https://doi.org/10.1016/j.jhlste.2021.100362>
- Nebel, S., Schneider, S., Beege, M., & Rey, G. D. (2017). Leaderboards within educational videogames: The impact of difficulty, effort and gameplay. *Computers and Education*, 113, 28–41. <https://doi.org/10.1016/j.compedu.2017.05.011>
- Nieto-Escamez, F. A., & Roldán-Tapia, M. D. (2021). Gamification as online teaching strategy during COVID-19: A mini-review. *Frontiers in Psychology*, 12, 648552. <https://doi.org/10.3389/fpsyg.2021.648552>
- Rincon-Flores, E. G., & Santos-Guevara, B. N. (2021). Gamification during Covid-19: Promoting active learning and motivation in higher education. *Australasian Journal of Educational Technology*, 37(5), 43–60. <https://doi.org/10.14742/ajet.7157>
- Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, 69, 371–380. <https://doi.org/10.1016/j.chb.2016.12.033>
- Sánchez-Mena, A., & Martí-Parreño, J. (2017). Drivers and barriers to adopting gamification: Teachers' perspectives. *Electronic Journal of E-Learning*, 15(5), 434–443
- Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human Computer Studies*, 74, 14–31. <https://doi.org/10.1016/j.ijhcs.2014.09.006>
- Shin, T. S., & Dickson, W. P. (2010). The effects of peer- and self-referenced feedback on students' motivation and academic performance in online learning environments. *Journal of Online Learning and Teaching*, 6(1), 187–197
- Tondello, G. F., Premasukh, H., & Nacke, L. E. (2018). A theory of gamification principles through goal-setting theory. In Proceedings of the 51st Hawaii international conference on system sciences, Hawaii, USA (pp. 1118–1127). <http://hdl.handle.net/10125/50027>. Accessed 30 May 2020.
- Wan, A. T., & Sadiq, S. (2012). Impact of personalized recommendation and social comparison on learning behaviors and outcomes. In *Proceedings of the 23rd Australasian conference on information systems, Geelong, Australia* (pp. 1–10)
- Wiersma, W., & Jurs, S. G. (2009). *Research methods in education: An introduction*. Allyn & Bacon
- Zainuddin, Z., Farida, R., Keumala, C. M., Kurniawan, R., & Iskandar, H. (2021). Synchronous online flip learning with formative gamification quiz: Instruction during COVID-19. *Interactive Technology and Smart Education*. Advance online publication. <https://doi.org/10.1108/ITSE-01-2021-0002>
- Zimmerman, B. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. <https://doi.org/10.3102/0002831207312909>
- Zimmerman, B. J., & Schunk, D. H. (2008). Motivation: An essential dimension of self- regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and Self- regulated Learning* (pp. 1–30). Erlbaum

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.