Computers & Education 144 (2020) 103666

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Gamification in the classroom: Examining the impact of gamified quizzes on student learning

Diana R. Sanchez[a,*,1], Markus Langer[b,1], Rupinder Kaur[a]

[a] San Francisco State University, USA
[b] Saarland University, Germany

ARTICLE INFO

Keywords:
Adult learning
Improving classroom teaching
Media in education
Distance education and telelearning

ABSTRACT

Gamification is emerging as a method aimed at enhancing instructional contents in educational settings. However, theoretical underpinnings of the proposed effects of gamification are lacking. This paper applies the theory of gamified learning and extends research exploring the benefits of gamification on student learning through the testing effect. In a quasi-experimental design, university students (N = 473) prepared for three tests using traditional quizzes (i.e., a question, four response options) or gamified online quizzes (i.e., a wager option, a progress bar, encouraging messages). We assumed that students completing gamified quizzes would complete more quizzes and, through the benefits of the testing effect, would demonstrate better learning. Findings supported the testing effect in that students who completed more quizzes performed better on subsequent tests. Furthermore, students who completed the gamified quizzes had significantly better scores on the first test. However, this effect was not due to students completing more quizzes in the gamification group. Additionally, the beneficial effect of gamification did not persist for subsequent tests. This supports that gamification might work through a novelty effect where its influence may not be sustainable. Further analyses showed that higher achieving students benefited more from gamification than lower achieving students. Overall, the results (a) imply that gamification may be a viable option for short-term assignments, (b) highlight concerns of a novelty effect possibly recommending instructors not to use the same gamification method permanently, and (c) indicate that there are contexts where gamification might not be adequate to target low achieving students. Given these results we call for longitudinal studies investigating the novelty effects of gamification and research examining individual differences moderating the effects of gamification.

The last decades have been hallmarked by rapid growth in technological development and innovation (Chang & Taylor, 2016). The education industry has capitalized on this trend by integrating new methods such as virtual collaboration, mobile learning applications, and other technology enhanced learning programs (Buabeng-Andoh, 2012; Domingo & Garganté, 2016; Irving et al., 2016; Mayer et al., 2009). Substantial research has explored these alternative classroom processes that aim to improve student-learning outcomes (Duchi, Hazan, & Singer, 2011; Garrison & Cleveland-Innes, 2005; Herrington, Oliver, & Reeves, 2003; Means, 2010; Means, Toyama, Murphy, Bakia, & Jones, 2009; Richardson & Swan, 2003; Sandholtz, 1997). This influx of research has included growing interest in alternative classroom structures such as augmented classrooms and blended learning (Barkley, Cross, & Major, 2014; Bonk & Graham, 2012; Kim, Kim, Khera, & Getman, 2014). However, restructuring classes is not always possible due to

* Corresponding author.
E-mail address: SanchezDianaR@sfsu.edu (D.R. Sanchez).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.compedu.2019.103666
Received 3 October 2018; Received in revised form 10 August 2019; Accepted 12 August 2019
Available online 13 August 2019
0360-1315/

logistical issues and limited resources. Educators therefore often search for alternative changes that can be easily implemented in order to improve learning. Gamification is one of these changes that may present itself as a useful, cost-effective, and efficient tool for educators to improve learning outcomes (Oprescu, Jones, & Katsikitis, 2014; Rowland, 2014).

Gamification refers to the use of game-design elements (e.g., points) and game characteristics (e.g., assessment, challenge) (for an overview see e.g. Bedwell, Pavlas, Heyne, Lazzara, & Salas, 2012) in non-game contexts in an attempt to achieve positive outcomes (e.g., enhance student learning) (Deterding, Khaled, Nacke, & Dixon, 2011). Given the implicit belief that games are enjoyable (Von Ahn & Dabbish, 2008), many instructors have integrated gamification into the classroom and researchers have studied the impact of gamification on classroom learning (e.g., Boticki, Baksa, Seow, & Looi, 2015; Hamari et al., 2016; Mekler, Brühlmann, Tuch, & Opwis, 2017).

Despite the widespread application and burgeoning research on gamification, the effects of gamification, its theoretical and psychological underpinnings, and individual differences that may affect gamification still lack understanding (Landers, Auer, Collmus, & Armstrong, 2018; Landers & Callan, 2011). To overcome this limitation, Landers (2014) introduced the theory of gamified learning in an attempt to provide research on learning through gamification with a theoretical foundation. He proposed that gamification can only benefit learning indirectly by improving already beneficial instructional content. Yet, there is still a lack of studies using this theoretical framework (or other theoretical frameworks, see e.g., Landers & Armstrong, 2017) when building a theoretical foundation on how gamification should affect learning (Landers et al., 2018; see Landers & Landers, 2014 as an exception). Furthermore, there is growing evidence that individual differences such as gender (Koivisto & Hamari, 2014) or video game experience (Landers & Armstrong, 2017) can affect benefits from gamification, calling for research exploring these moderating effects, especially in educational settings.

Therefore, the current study aims to enhance the understanding of the circumstances under which gamification can improve learning outcomes (e.g., Koivisto & Hamari, 2014; Landers, Bauer, & Callan, 2017; Sailer, Hense, Mayr, & Mandl, 2017) by building on the theory of gamified learning (Landers, 2014) and the well-established educational benefits of the testing effect (i.e. the effect that testing can enhance learning through the cognitive process of recalling information; Baird, 1985; Rowland, 2014). Specifically, following the propositions of the theory of gamified learning, gamification (i.e., the addition of game elements, in our case a wager option, a progress bar, and encouraging messages, all adding the game characteristics assessment and goals/rules from Bedwell et al., 2012 to the quizzes) may have the potential to enhance the testing effect. We postulate that students might engage in more preparational quizzes when game elements are added (cf., Landers & Landers, 2014). Furthermore, we investigate the potential moderating effect of one of the most important individual differences in the educational context: student achievement (Richardson, Abraham, & Bond, 2012).

Thus, the goals of this study are to (a) apply the theory of gamified learning as a theoretical model for the assumed effects of gamification, (b) replicate the testing effect in a web-based learning environment, (c) test if gamification can enhance the testing effect, and (d) determine if there are student characteristics that affect the effects of gamification. Therefore, the current study aims to contribute to research showing the value of the theory of gamified learning for developing theoretical models for the effects of gamification (see also Landers & Landers, 2014). This study also provides instructors with insights on the potential effects of gamification within the setting of online quizzes and shows that instructors need to be aware of student characteristics that may moderate the effects of gamification.

1. Theoretical background

1.1. Technology and learning

The integration of technology into our daily lives has reached the classroom as instructors utilize new technological resources to aid classroom instruction (Green & Hannon, 2007). Students can work collaboratively on online projects, discuss lectures using online forums, watch videos embedded into an instructor's lecture notes (Narciss, Proske, & Koerndle, 2007), or use a combination of the aforementioned approaches within social network sites developed for learning purposes (Landers & Callan, 2011). Technology can enhance learning through several processes such as providing students with instant feedback, making additional resources readily available, or allowing them to practice their skills at their own pace and test their own knowledge (Sitzmann, Kraiger, Stewart, & Wisher, 2006).

This study explores the use of technology through quizzes presented to students in an online learning management system. In these self-paced online quizzes, students can test their knowledge on the respective learning module and receive feedback about their performance. We compare two courses that used these online quizzes, where students in one of the courses received a gamified version of these quizzes. We propose that gamification of these quizzes will lead to better learning outcomes. In order to develop a theoretical model that underlies the effects of the instructional design within the current study, we build upon the theory of gamified learning (Landers, 2014). Following, we will introduce the theory of gamified learning and ground our assumptions and hypotheses on its propositions. Specifically, we argue that (a) completing online quizzes will lead to better learning outcomes, (b) students will complete more online quizzes when the quizzes are gamified which will (c) consequently lead to better learning outcomes. Furthermore, we will examine if student achievement, known to be related to students' effort regulation (Richardson et al., 2012), will affect the benefits from gamification. Fig. 1 shows the underlying theoretical model of this study, which is based on the theory of gamified learning and which includes its propositions and our hypotheses that will be tested during this study.

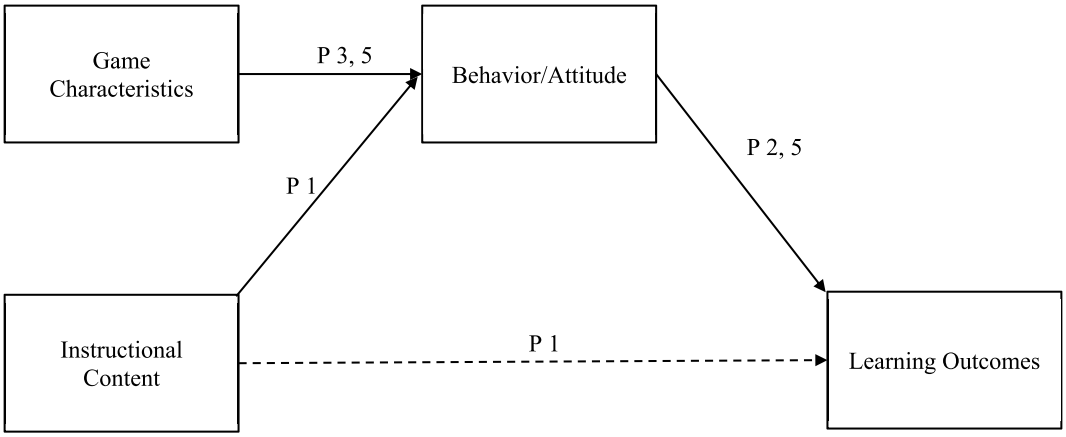

Fig. 1. The theory of gamified learning by Landers (2014), p. 760. Note that the path Game Characteristics → Behavior/Attitudes → Learning Outcomes, as well as the path Instructional Content → Behavior/Attitude → Learning Outcomes are mediating processes. The original model by Landers (2014) also included a moderating process (original proposition 4) from Behavior/Attitude on the path Instructional Content → Learning Outcomes. We left out this path as we did not hypothesize a moderating process resulting from our Behavior/Attitude measure (i.e., number of Quizzes Completed).

1.2. The theory of gamified learning

Gamification is defined as the integration of game elements into non-game environments (Dicheva, Dichev, Agre, & Angelova, 2015). For instance, Landers and Callan (2011) developed a social network site for student learning where students could discuss learning issues and share learning experiences. Especially important for the current study, they introduced gamified online quizzes in an attempt to encourage students to learn the instructional contents from these quizzes. Students who completed these optional quizzes could receive badges and level-ups when successfully completing quizzes (i.e., badges and leveling-up were the game elements). As a result of their study, they found that most students enjoyed the additional learning opportunity through gamified quizzes.

The study by Landers and Callan is an example of how to use technology and gamification within learning contexts and is a call for structured attempts to investigate the impact of gamification on student learning. For instance, they assumed that gamification would encourage students complete the online quizzes. However, they did not compare the gamified version of the quizzes to a non-gamified version, thus it remained unclear if gamification led to higher student engagement or to actual learning benefits.

Three years after the study by Landers and Callan, Landers (2014) introduced the theory of gamified learning as an important milestone, aiding subsequent studies approach research on gamification in a structured manner (see for instance Landers & Landers, 2014 investigation of the influence leaderboards have on student engagement and testing of the theory of gamified learning). The core idea behind this theory is that gamification can improve existing instructional content either through a moderation or a mediation process. There are five propositions of the theory of gamified learning, four of which have to be met in cases where gamification influences learning outcomes through a mediating process as we expect in the current study[2]: "Instructional content influences learning outcomes and behaviors (original Proposition 1); behaviors/attitudes influence learning (original Proposition 2); game characteristics influence changes in behavior/attitudes (original Proposition 3); the relationship between game elements and learning outcomes is mediated by behaviors/attitudes (original Proposition 5)" (Landers, 2014, pp. 760–762; the theory of gamified learning and these propositions are displayed in Fig. 1). Following, we will first explain the terms from the theory of gamified learning in the context of the current study, introduce the aforementioned propositions, and connect them to the hypotheses of the current study.

1.2.1. Clarification of terms

The theory of gamified learning includes four main terms: instructional content, game characteristics, behavior/attitudes, and learning outcomes. First, instructional content is defined as the instructions instructors use to educate students and facilitate student understanding (Reigeluth, 1983). In the current study, online quizzes constitute the instructional content that is meant to help students improve their declarative knowledge. Second, game characteristics are the game elements used to gamify the instructional content. Bedwell et al. (2012) introduced a framework of nine game characteristics that organizes game elements into categories. Using this taxonomy, the current study applies the game characteristics rules/goals (i.e., allowing students to wager points for a hint on the correct answer to a quiz question) and assessment (i.e., by showing students a progress bar of earned and possible points for a

[2] Note, that we left out original proposition 4 of the theory of gamified learning (Landers, 2014) which states that "Game elements affect behaviors/attitudes that moderate instructional effectiveness" (Landers, 2014, p. 761) as we do not postulate a moderating effect of the game elements on instructional effectiveness.

quiz, and by providing an encouraging message with their final score). Third, behavior/attitude refers to behaviors and attitudes affected by the instructional content and the game characteristics. In the current study, the instructional content and game characteristics will most likely affect completing online quizzes (behavior) reflecting student engagement with the quizzes (attitude) (cf., Landers & Landers, 2014). Finally, the main outcome in the theory of gamified learning are learning outcomes. In the case of our study, students will learn the content of the course which will be demonstrated by higher scores in three declarative knowledge tests given throughout the course.

1.2.2. Quizzes affect learning

In the first proposition of the theory of gamified learning, Landers stated that if gamification aims to be successful it must be certain that the "instructional content influences learning outcomes and behaviors" (Landers, 2014, p. 760). In other words, if students do not benefit from the instructional content, there is no benefit in gamifying it. Translated to the current study, this means it is a precondition of potential benefits of gamification that providing quizzes already aids student learning.

The effect that quizzes and tests can improve student learning is known as the testing effect (see Rowland, 2014 for a meta-analysis). Similar terms include "retrieval practice" or "test-enhanced learning", which refer to the increased retention of learned knowledge and/or skills by retrieving material through testing (Larsen, 2013). Numerous studies have been conducted to examine and confirm the testing effect in experimental settings (e.g. Jensen, McDaniel, Woodard, & Kummer, 2014; Johnson & Mayer, 2009; McDaniel, Agarwal, Huelser, McDermott, & Roediger, 2011; McDaniel, Anderson, Derbish, & Morrisette, 2007). Educators have also applied this concept to the classroom to determine if various practices can improve student learning (Carpenter, 2012; Rowland, 2014; Vojdanoska, Cranney, & Newell, 2010). The literature predominantly supports that the testing effect benefits students by increasing learning and retention (Bangert-Drowns, Kulik, & Kulik, 1991; McDaniel, Roediger, & McDermott, 2007; Rowland, 2014). Its effectiveness is attributed to more effortful retrieval (e.g., compared to repeatedly reading through learning materials; Rowland, 2014), the strengthening of neural pathways, and increasing the number of neurological connections, which makes the recalled information more accessible (Carpenter & DeLosh, 2006). For instance, Wing, Marsh, and Cabeza (2013) examined brain activities when utilizing the testing effect and found that testing contributed to future memory success and was associated with different brain activity in comparison to traditional studying of learning material. An fMRI analysis revealed activation of brain areas, such as the anterior hippocampus and lateral temporal cortices, suggesting that the testing effect uniquely activates certain neural pathways. Additionally, Wheeler and Roediger (1992) found improved retention of information after repeated retrieval of that information in comparison to little or no retrieval. These results demonstrate that increased retrieval of information leads to deeper processing and improved retention results, which constitutes the testing effect.

Regarding proposition 1 of the theory of gamified learning, previous research supports that quizzes can be effective as instructional content. In other words, we assume that the guidelines of proposition 1 are met within our research model because the instructional content of providing students with quizzes builds on the strongly supported testing effect.

1.2.3. More quizzes improve learning outcomes

The second proposition from the theory of gamified learning by Landers is that "behaviors/attitudes influence learning" (Landers, 2014, p. 760). In the case of the current study, students had the opportunity to complete online quizzes in exchange for course credit and to deepen their understanding of course content. According to previous research, students completing (more) quizzes will have improved learning which will lead to better performance in subsequent tests (Kling, McCorkle, Miller, & Reardon, 2005; McDaniel, Anderson, et al., 2007, McDaniel, Roediger, et al., 2007; McDaniel et al., 2011). This means that completing online quizzes affects learning outcomes positively, thus students who complete more quizzes may show better learning outcomes. The current study will replicate the testing effect through online quizzes to (a) increase confidence that our sample is representative of the general population (b) support the testing effect in technology-based settings, and (c) support proposition 2 of the theory of gamified learning within our research model. Therefore, our first hypothesis is as follows:

Hypothesis 1. There will be a significant relationship between the number of quizzes students completed and students' test scores (for Test A, Test B, and Test C).

1.2.4. Gamification leads to completing more quizzes

Proposition 3 from the theory of gamified learning by Landers is that "game characteristics influence changes in behavior/attitudes" (Landers, 2014, p. 761). While the term "gamification" has relatively new applications in the research literature, the idea of using game mechanics to engage students and other audiences has a long-standing history (e.g. Harvey, 1970). This strategy has been used in a number of applications by the military, philosophers, and researchers (Zichermann & Cunningham, 2011). Gamification is used in many contexts, such as survey (Wells, 2016) and mobile applications development (Lister, West, Cannon, Sax, & Brodegard, 2014). It has also been used in workplace contexts (Singh, 2012), for example to prepare applicants for job interviews (e.g., Gebhard et al., 2018; Langer, König, Gebhard, & André, 2016), for employee selection (Armstrong, Landers, & Collmus, 2016; Oprescu et al., 2014), and to train employees (Cornelissen et al., 2013; Di Bitonto, Corriero, Pesare, Rossano, & Roselli, 2014). Most important for the current study is the application of gamification to train working memory (Farcas, Szamosközi, & Takacs, 2016), within online educational courses (Chang & Wei, 2016), in order to engage students (Landers & Landers, 2014), and in further educational environments (Canhoto & Murphy, 2016) supporting the potential value and practical relevance of gamification for learning. We focus on the possible effects of gamification on student engagement as this is most likely to relate to the benefits of gamification on the testing effect.

1.2.5. Student engagement

Because there are variations in the definition and application of the term engagement, we clarify for this study that engagement is students' allocation of resources and energy to the quizzes (Rich, Lepine, & Crawford, 2010). Note that during data collection there was no opportunity to directly observe or measure student engagement during the quizzes, therefore the current study uses voluntary task re-engagement as an indicator of engagement (cf., Ryan, Koestner, & Deci, 1991). Previous definitions and studies have interpreted voluntarily re-engagement in an activity without specific instructions during a period of free-choice to demonstrate intrinsic motivation and engagement (Ryan et al., 1991). We assume that students who complete more quizzes voluntarily are more engaged in the task because they will choose to re-engage in the task without specific instructions to do so.

Gamification is often aimed at benefiting learning by increasing engagement in a task (Brull & Finlayson, 2016). For instance, Landers and Landers (2014) used leaderboards to engage students to work on an online wiki. They found that through leaderboards which, similar to the gamification elements of the current study, relied on the game characteristics assessment (students could see how many tasks they fulfilled) and rules/goals (students could see which tasks they need to fulfill to receive points to climb the leaderboard) students spent more time on developing the online wiki than in a control group without a leaderboard. In the current study, the gamified version of the online quizzes includes game elements based on the same game characteristics (i.e., assessment and rules/goals). Specifically, there are three game elements within this study which focus on the game characteristics assessment and rules/goals that have engaging potential: a progress bar, encouraging messages, and the possibility to wager points for a hint.

The first game element included in the quizzes is a progress bar. Potential benefits and gamification aspects of progress bars have been discussed in previous research (Dicheva et al., 2015; Zichermann & Cunningham, 2011). The progress bar in this study displayed students' current points as a dark shaded section and potential points as a lightly shaded section of what they could earn if they answered the current question correctly. Progress bars show individuals how far they have already come and creating an incentive to reach a 100% completion (i.e., a full progress bar). This falls under the game characteristics assessment and rules/goals as defined by Bedwell et al. (2012) because it is providing feedback to the individual to assess their own performance and presents them their goal explicitly (i.e., a full progress bar reflects the goal). The progress bar might engage students as it generates a sense of closeness to the points and the goal (cf., Dicheva, Irwin, Dichev, & Talasila, 2014) as well as a mastery experience that might grow the more students proceed within the progress bar (Dicheva et al., 2015).

The second game element is an encouraging message displayed at the end of the quiz. This feature urges students who earned partial credit to try again for full credit, congratulates students who earned full credit, and reminds them that they can continue taking the quiz to prepare for the upcoming test. Encouraging messages correspond to the game characteristic assessment. They give students feedback while also reminding them of the benefits of returning to the quiz. Similar encouraging messages have been addressed in previous research. For example, Zichermann and Cunningham (2011) discuss games such as *QuantaPet* which encourages players to return to the game through incentivized behavior such as taking care of digital bets. Encouraging messages are also often added to mobile applications because they are believed to keep users returning to the application. For instance, the application “Duolingo” is intended to improve foreign language skills and promote users returning to the app with several gamification techniques (e.g., points, badges for achievements, content unlocking). This specific app also uses encouraging messages to congratulate learners and to remind them to keep up the good work. This feature might engage learners by appealing to their need for achievement and furthermore provides positive feedback on the effort students make to take a quiz. Based on previous research, there is an understanding that positive feedback and encouraging messages can lead to positive experiences, motivation (Koka & Hein, 2003) and improved learning in an activity (Goh, Seet, & Chen, 2012). Additionally, by encouraging students to take the quiz again, it should appeal to student's desire to reach a score of 100% or be prepared for the next test.

The third game element is a wager option where students could select an icon of a magnifying glass labeled as hint. They would wager ½ point to reduce the response options from four to two, giving them a 50/50 chance of selecting the correct answer. Research on betting and wagering usually explores the less savory aspects of addiction and gambling (e.g., Gray, LaPlante, & Shaffer, 2012). However, educators may acknowledge the addictive properties of gamification as a desirable feature they would wish to capture and utilize to engage students towards returning to the learning experience (Garris, Ahlers, & Driskell, 2002). Although it has been discussed briefly in other articles (e.g., Coccoli, Iacono, & Vercelli, 2015), little research has been done on how wagering and betting features impact learning through gamified experiences. The wager option is part of the rules/goals characteristic that tells individuals about the parameters within the learning activity that reward them with points. In addition, it also gives the player more control of the activity and the potential outcome. Research has shown that individuals carefully consider the perceived risks and benefits of a situation when presented with these options. Given that the hint option would reduce the risk of selecting an incorrect answer but also reduce the benefit through only receiving half a point for a correct answer, students would need to weigh these options carefully (Kahneman & Tversky, 1984). In order to achieve the highest scores, students' goal should be to try to find the correct answer independently of the hints. In the first quizzes, students might still use the hints but with every new try, students who want to receive high scores should be motivated to not use the hints any more. The process of maturing from using these hints may encourage students to complete the quizzes several times and it could reflect learning progress, as well as an engaging experience as students realize that they do not need the hints any more. In sum, including the wager option could encourage the students to complete more quizzes because completing the gamified quizzes may lead to (a) motivation to beat their own high score, (b) a sense of achievement, and (c) a better perception of the learning progress.

Overall, providing students with encouraging messages provides positive feedback and should encourage students to return to the task to improve their results. Visually tracking points with the progress bar may create a sense of closeness to their goals as well as a

mastery experience on their way to the learning goal (cf., Dicheva et al., 2015). Allowing students to wager points for hints could create a sense of control as well as encourage students to improve by not needing the hint option any more. These may all create a gamified experience for students and engage them to complete the quizzes.

1.2.6. Gamification affects learning outcomes through completing quizzes

The last proposition from the theory of gamified learning by Landers is that "the relationship between game elements and learning outcomes is mediated by behaviors/attitudes" (Landers, 2014, p. 762). One primary purpose of this study is to understand if gamification can enhance the benefits of the testing effect. Our intention is to expand existing literature on a media enhanced testing effect (e.g., Johnson & Mayer, 2009) and to explore a specific type of media enhancement – gamification. There is reason to believe that the concept of gamification could positively increase the impact of the testing effect as gamification is believed to function through factors such as student engagement (Su & Cheng, 2015). According to Lander's theory of gamified learning (2014) and results from Landers and Landers (2014), gamification can affect students' motivation to invest in learning. In the current study, this could mean that gamifying quizzes will lead to more student engagement when completing the quizzes (i.e., students will take more quizzes when the quiz is gamified; which would support proposition 3 from the theory of gamified learning) which may lead to better learning outcomes (i.e., performance in tests on the same learning content). In other words, we predict that gamification leads to students being more motivated and engaged to fulfill online quizzes which consequently leads to better learning through the learning mechanisms of the testing effect (which would support proposition 5 of the theory of gamified learning) (cf., Kling et al., 2005; McDaniel, Anderson, et al., 2007, McDaniel, Roediger, et al., 2007; McDaniel et al., 2011). We thus propose:

Hypotheses 2. The positive effect of gamification on test scores (Test A, B, and C) will be mediated by the number of Quizzes Completed.

1.3. Student abilities moderating the effects of gamification

As a further step of our data analysis we investigated the moderated-mediation model displayed in the model in Fig. 2. The

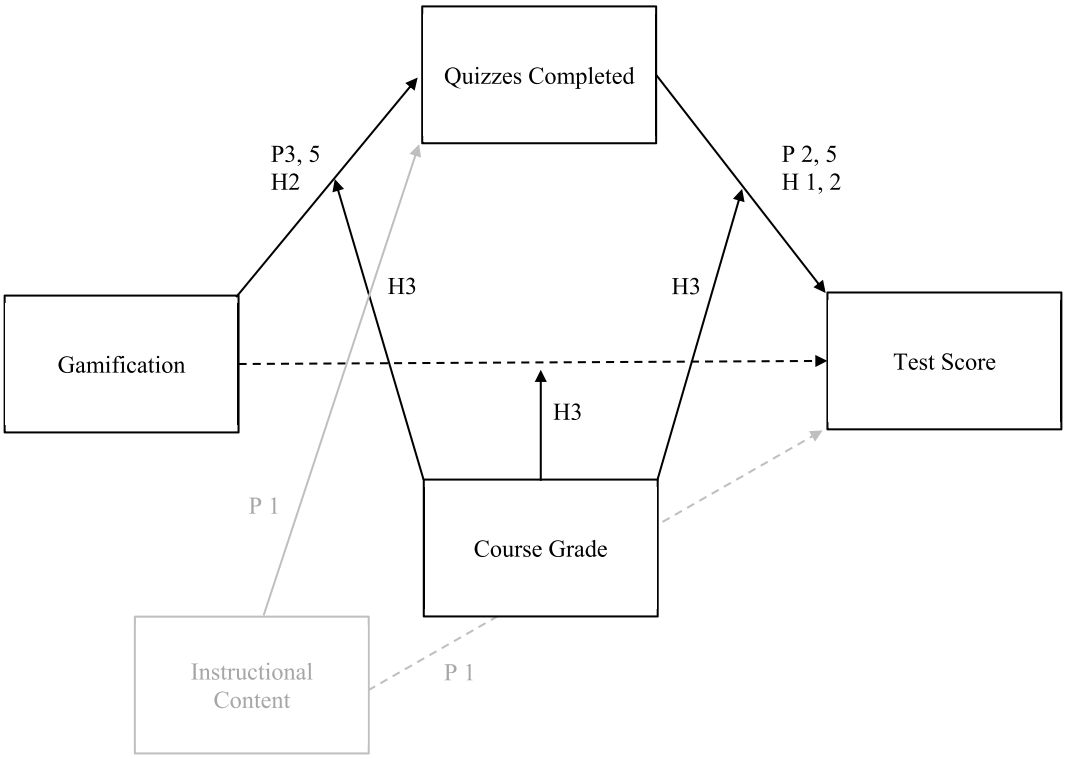

Fig. 2. Theoretical model of the current study based on the theory of gamified learning (see Landers, 2014, p. 760) with the addition of course grade as a moderator and the exclusion of the Instructional Content box from the original theoretical model (see grey box and paths). In order to match the wording within our study, we adapted the content of the boxes (i.e., Learning Outcomes from the original theory of gamified learning are the Test Scores within the current study; Game Elements from the original theory is Gamification; Behavior/Attitudes from the original theory are Quizzes Completed). The path Gamification → Quizzes Completed → Test Score displays a mediation process. The influence of Course Grade on the paths Gamification → Quizzes Completed and Quizzes Completed → Test Score displays moderating processes. P = Proposition from the theory of gamified learning. H = Hypothesis in the current study.

mediation part of these models is based on the potential effect of the gamified quiz format on the number of Quizzes Completed (i.e., Hypothesis 2). Furthermore, it might be possible that individual differences affect the influences of gamification on student behavior or learning outcomes. For instance, Orvis, Horn, and Belanich (2009) found that game self-efficacy and goal orientation affected learning from a game and Landers et al. (2017) investigated the role of goal-commitment when using the game element leaderboards to affect task performance.

In educational contexts, one of the most important individual differences that may affect the outcomes of gamification is students' ability to study, which previous research commonly measured using students' academic achievement (Richardson et al., 2012; Zimmerman, Bandura, & Martinez-Pons, 1992). In the current study, academic achievement is reflected by students' overall course grade. Richardson et al. (2012) showed in their meta-analysis that students' academic achievement relates to their academic self-efficacy and effort regulation. The latter variable might especially influence the assumed effects of our gamification approach on students as it is closely tied to students' motivation and persistence in academic tasks such as seriously investing in given instructional content (e.g. online quizzes) (Richardson et al., 2012). Therefore, students' abilities may affect the relationships between the gamified quiz format and test scores, gamification and the number of Quizzes Completed, as well as the number of Quizzes Completed and test scores. The moderation in Fig. 2 investigates the aforementioned assumptions. First, it may be that higher achieving students are more (or less) engaged by the gamified quiz format and in the end complete more (or less) quizzes than lower achieving students who may be less (or more) affected by the gamified quiz format. Second, it could be that higher achieving students benefit more (or less) from completing more quizzes, which would affect the relationship between the number of Quizzes Completed and the test scores. Third, it might also be that the gamified quiz format affects higher achieving students test scores more directly or through variables that were not captured in this study, which would affect the direct effect of the quiz format on test scores.

Exploratory Hypothesis 3. : Student's course grades may moderate the relationships between the quiz format and the number of Quizzes Completed, between number of Quizzes Completed and test scores, and between the quiz format and test scores.

2. Method

2.1. Participants

Information was collected from an archival dataset from the psychology department at a Western University. The dataset contained information from two consecutive semesters of students enrolled in an introductory psychology course. Due to the limitations of using an archival dataset, demographic information was limited to gender, year in school, and major. The total sample consisted of 473 students (i.e., 316 completed the traditional quiz and 157 completed the gamified quiz). This data was collected using archival resources across classes of students taking the same course. Using archival data made random assignment to the two Quiz Formats impossible, which is why this study is based on a quasi-experimental design. Thus, students in the same class were assigned to the same Quiz Format (i.e., data using the traditional quiz was collected from two classes of students and data for the gamified quiz was collected from the third class of students). This resulted in an uneven distribution of participants in the two conditions, with two thirds of the sample assigned to the traditional quiz. Students represented over 50 different majors, the majority reporting an undeclared major 22%, followed by 13% in Health and Exercise, 10% in Business Administration, and 6% in Biological Science. In the traditional quiz group, there were 51.3% males compared to 54.8% males in the gamified quiz group which constitutes no significant difference in the gender distribution, $\chi^2(1) = 0.51$, $p = .47$. Regarding year at the college, in the traditional quiz group, there were 21.8% freshman, 50.3% sophomores, 16.8% juniors, and 11.1% seniors compared to 63.1% freshman, 22.9% sophomores, 9.6% juniors and 7% seniors in the gamified group which is a significantly different distribution regarding year at the college, $\chi^2(1) = 78.24$, $p = .01$.

2.2. Procedures

The course included three tests across the semester (i.e., Test A, Test B, and Test C). Note that previous research implied that there might be a novelty effect when using gamification (Hamari, Koivisto, & Sarsa, 2014). Specifically, there are previous findings where the engaging and beneficial effects of gamification dissipate over time (Koivisto & Hamari, 2014). The longitudinal design of our data (i.e., a series of three tests students take over a semester) allows to observe if and at what point in time the effects of gamification diminish over time which could indicate a novelty effect. Each test covered material from multiple chapters of content. For each content chapter, students first read the chapter and could complete online quizzes on the content. Students then received a lecture on the content. Both semesters were identical (i.e., with the same instructor, textbook, lectures, resources, and tests). The only controlled difference between the two semesters was the format of the online quizzes that was given to the students, described below. At the end of the semester, student information, quiz scores, test scores, and final course grades were collected.

2.3. Online quizzes

During the semester, students were invited to complete a series of online quizzes (in total 34 quizzes) each containing five multiple-choice questions drawn from a pool of questions on the respective topic. The number of multiple-choice questions in each pool varied with the total number of 775 multiple-choice questions across all the topics. Quizzes for both experimental conditions used the same question prompt and four response options. There was only one correct answer for each question and students were

given instructions that the quizzes were untimed and open-book. After submitting a response, students were given feedback on whether or not the response was correct and what the correct answer was before moving on to the next question. Course credit was awarded to students for completing the quizzes. This was intended to incentivize students to complete the quizzes. The points earned in the quiz directly translated to points in the class, which were applied towards the final grade. Because each quiz had five questions, students could earn up to five points for each quiz if they answered all of the questions correctly. If a student only answered four of the five questions correctly, they only earned four points towards their grade. Quizzes could be taken an unlimited number of times; however, only the highest score a student earned across all attempts of a single quiz was recorded towards their final grade. Therefore, even if a student earned full credit of five points right away, there was still an incentive to continue taking the quizzes to help prepare for the upcoming tests.

2.4. Quiz format

The two quizzes only differed in their format. The traditional quiz presented the question, the response options, and a next button. After each question feedback was given. At the end of the quiz a final score was provided, showing how many points the student earned out of the possible five points. The gamified quiz format added three game elements. The first was a progress bar. At the beginning of each question a faded section would be added to the progress bar displaying the *potential points* a student could earn for answering the question correctly. If the student answered the question correctly, the color would fade-in, demonstrating they earned the points for that question. With an incorrect answer the *potential points* faded-away from the progress bar. The second element of the gamified quizzes was a wager option. In the gamified quiz students saw the image of a magnifying glass in the top right corner with the word “Hint” next to it. If students clicked this button the system would ask them if they would like a hint; offering to take away two of the possible response options for reduced points. If students accepted the hint, two of the four response options were removed and the section of the faded progress bar reduced, decreasing the worth of the question to ½ of the points if answered correctly. The third feature of the gamified quizzes was an encouraging message that students received at the end of the quiz. Upon completion of the last question, students saw one of two messages. If student answered one or more of the questions incorrectly, the message encouraged them to try the quiz again to try and reach 100%. If the student answered all five questions correctly, the message congratulated them and encouraged them to return to the quiz to keep practicing for the upcoming test.

2.5. Measures

2.5.1. Test scores

Students were given three tests during the course. Each test was given across the 16-week semester. Test A was given during week five, Text B during week eight and Test C during week 13. Each test consisted of 55 multiple-choice questions with four response options and one correct answer. All of the test questions were selected from a large pool of questions using item statistics (Findley, Keever, Chappelka, Eakes, & Gilliam, 1996) gathered from a previous semester. Test scores are used as a measure for student learning since students who learned more from the quizzes would perform better in these tests.

2.5.2. Number of quizzes completed

Students had the opportunity to complete quizzes on each topic area for the course. Every time a student completed a quiz, this increased the number of Quizzes Completed for this student. The total number of Quizzes Completed for each student were summed up separately for quizzes before Test A, B, and C.

2.5.3. Course grade

All students received a cumulative grade for the class at the end of the course which encapsulated their performance across the semester. The Course Grade included scores from students for other assignments in the class (e.g., research papers, thought papers). Scores from Test A, Test B, and Test C, as well as scores for all of the quizzes, were removed from the calculation of the Course Grade. Course Grade should therefore be a proxy measure for student achievement and their general capabilities to study at a university.

2.6. Additional analyses

In order to test the moderated-mediation model in Fig. 2, we used the PROCESS macro for SPSS (Hayes, 2015). For every test, we included the respective number of Quizzes Completed (e.g., number of Quizzes Completed related to Test A) as the mediator and the students' Course Grade as the moderator for the relation between Quiz Format and number of Quizzes Completed, number of Quizzes Completed and Test Scores (i.e., Test A, Test B, and Test C), and for the direct effect of the Quiz Format on Test Scores. The outcome variable was the respective Test Score. With PROCESS it is possible to evaluate mediation and moderation effects step-wise (for a detailed description see Hayes, 2015). First, PROCESS provides an output for the effect of the independent variable onto the mediator variable including the moderator effect displayed as the interaction between the independent variable and the moderator (in our case Quiz Format x Course Grade). Second, it offers an output indicating whether the mediating variable impacts the outcome under the condition that the independent variable is also included in the regression model. In this overall model it includes the interaction effects of the mediator and the moderator (in our case number of Quizzes Completed x Course Grade) and the interaction of the independent variable and the moderator. Third, PROCESS provides bias-corrected bootstrapped estimates of the confidence intervals for the conditional direct effect. In the given case, this means that it calculates the direct effect of the Quiz Format on the Test Scores

Table 1
Descriptives and correlations of study variables.

	n	*M*	*SD*	1	2	3	4	5	6
1. Gender	473	–	–						
2. Quiz Format	473	–	–	.03					
3. # of Quizzes Completed	473	18.15	5.94	.00	-.03				
4. Test A	454	.79	.13	-.01	.09*	.26**			
5. Test B	444	.71	.14	.01	.05	.31**	.68**		
6. Test C	437	.72	.13	-.07	.09	.23**	.66**	.68**	
7. Course Grade	473	.79	.18	.06	-.21**	.23**	.24**	.29**	.22**

Note. Gender, female = 0, male = 1; Quiz Format, traditional quiz = 0, gamified quiz = 1. Test A, Test B, Test C, and Course Grade are all percentages.
*$p < .05$, **$p < .01$.

for students at the mean of the moderator Course Grade and one standard deviation above (higher achieving students) and below the mean (lower achieving students). Simultaneously, it also provides the conditional indirect effects. In our case these are the effects of the Quiz Format on Test Scores mediated by number of Quizzes Completed at the mean of the moderator Course Grade and one standard deviation above and below the mean. To be clear, if the confidence intervals do not include zero, this indicates significant direct or indirect effects of the independent variable on the outcome on a given level of the moderator.

3. Results

Correlations between the study variables as well as descriptives are provided in Table 1.

Hypothesis 1 proposed that there will be a significant relationship between the number of quizzes a student completed and the student's Test Scores. Our results support Hypothesis 1 showing that students who completed more quizzes performed significantly better than students who completed fewer quizzes; Test A, $F(1,453) = 3.85$, $p < .01$, *partial* $\eta^2 = 0.05$, Test B, $F(1,443) = 6.81$, $p < .01$, *partial* $\eta^2 = 0.09$, and Test C $F(1,436) = 1.94$, $p < .05$, *partial* $\eta^2 = 0.05$.

Hypothesis 2 stated that the positive effect of gamification on Test Scores will be mediated by the numbers of Quizzes Completed. As a first step to examine this hypothesis, we conducted ANOVAs to reveal if there were differences between the two Quiz Formats on the Test Scores. For Test A, students who completed the gamified quizzes ($M = 80.27$, $SD = 14.32$) had significantly higher scores compared to students who completed the traditional quizzes ($M = 77.69$, $SD = 12.19$), $F(1,453) = 4.03$, $p < .05$, *partial* $\eta^2 = 0.01$. However, the test scores were not significantly different for both remaining tests; Test B gamified quizzes ($M = 71.91$, $SD = 14.07$) and traditional quizzes ($M = 74.06$, $SD = 12.18$), $F(1,443) = 1.29$, $p = .26$, *partial* $\eta^2 = 0.003$ and Test C, gamified quizzes ($M = 70.31$, $SD = 13.84$) and traditional quizzes ($M = 71.66$, $SD = 13.05$), $F(1,436) = 3.40$, $p = .07$, *partial* $\eta^2 = 0.008$.

We continued to examine Hypothesis 2 using PROCESS. As this study is based on a quasi-experimental design, we dummy-coded participants' year of study and included it as a covariate in addition to participants' gender into the PROCESS analyses. These results differed from the model without any covariates only in numerical terms but not in terms of significance and interpretation of the results. For the sake of simplicity, we therefore present the results without any covariates.[3] Results of the moderated-mediation analyses testing the model in Fig. 1 are presented in Table 2, Table 3, and Fig. 3. In contrast to what we proposed in Hypothesis 2 (that gamification will positively affect Test Scores through students' taking more quizzes), results imply that gamification had a negative effect on the number of Quizzes Completed for Test A. This means that students who completed the gamified quizzes completed fewer quizzes in preparation of Test A than students assigned to the traditional quizzes. However, students who took the gamified quizzes outperformed the students in the other group regarding Test A. For the other tests there was neither a positive nor a negative effect of gamification on Test Scores through the number of Quizzes Completed. Therefore, there was no support for Hypothesis 2.

In exploratory Hypothesis 3 we proposed that student's Course Grades may moderate the relationship between the Quiz Format and the number of Quizzes Completed, number of Quizzes Completed and Test Scores, as well as the relation between the Quiz Format and Test Scores. The most robust result was found for the conditional direct effect of Quiz Format on Test Scores. Table 2 shows that for all analyses of the complete models (i.e., complete Model Test A, Test B and Test C), the interaction between the Quiz Format and the Course Grade is significant which is a first indicator for a moderation effect. This effect is further supported by Table 3 which shows that for Test A, B, and C, the direct positive effects of gamification on the Test Scores were only significant for students with a medium and a high Course Grade (indicated by the confidence intervals not including zero for students with medium and high Course Grades). This implies that Course Grade moderated the direct effect of the Quiz Format on the Test Scores. In other words, this indicates that students with lower Course Grades did not benefit directly from the gamification of the quizzes in regard of their Test Scores, whereas students with average and high Course Grades did. These results are displayed in Fig. 3.

There were no consistent findings regarding the interaction of the Quiz Format and the Course Grade in the models testing the effect of the Quiz Format on the number of Quizzes Completed. This means that Course Grade did not moderate the effect of the Quiz Format on the number of Quizzes Completed (i.e., gamification did not have a different effect on the number of Quizzes Completed

[3] The results including covariates can be made available on request.

Table 2
Regression results for the tests regarding the moderated-mediation model with the mediator Quizzes Completed and the moderator course grade (see Fig. 1).

Model	R^2	Coefficient	*SE*	95% Confidence Interval
Effect on Quizzes Completed A	.06**	–	–	–
Quiz Format → Quizzes Completed A		-.25**	.09	[-.42, −.07]
Course Grade → Quizzes Completed A		.13*	.06	[.01, .25]
Quiz Format x Course Grade → Quizzes Completed A		.08	.09	[-.10, .26]
Complete Model Test A	.17**	–	–	–
Quizzes Completed A → Score Test A		.21**	.05	[.11, .30]
Course Grade → Score Test A		.00	.07	[-.13, .13]
Quiz Format → Score Test A		.38**	.09	[.20, .56]
Quizzes Completed x Course Grade → Score Test A		.02	.04	[-.06, .10]
Quiz Format x Course Grade → Score Test A		.55**	.10	[.36, .73]
Effect on Quizzes Completed B	.06**	–	–	–
Quiz Format → Quizzes Completed B		-.06	.09	[-.23, .12]
Course Grade → Quizzes Completed B		.11	.08	[-.01, .22]
Quiz Format x Course Grade → Quizzes Completed B		.17*	.09	[.003, .34]
Complete Model Test B	.17**	–	–	–
Quizzes Completed B → Score Test B		.26**	.05	[.16, .37]
Course Grade → Score Test B		.13*	.07	[.01, .26]
Quiz Format → Score Test B		.27**	.10	[.09, .46]
Quizzes Completed x Course Grade → Score Test B		.04	.04	[-.04, .10]
Quiz Format x Course Grade → Score Test B		.30**	.09	[.11, .48]
Effect on Quizzes Completed C	.00	–	–	–
Quiz Format → Quizzes Completed C		.07	.09	[-.10, .24]
Course Grade → Quizzes Completed C		.05	.06	[-.08, .17]
Quiz Format x Course Grade → Quizzes Completed C		.05	.11	[-.16, .26]
Complete Model Test C	.13**	–	–	–
Quizzes Completed C → Score Test C		.19**	.05	[.08, .29]
Course Grade → Score Test C		.12	.07	[-.02, .26]
Quiz Format → Score Test C		.23*	.10	[.04, .42]
Quizzes Completed x Course Grade → Score Test C		.06	.06	[-.06, .17]
Quiz Format x Course Grade → Score Test C		.50**	.12	[.26, .74]

Note. Coding of the variable Quiz Format: 1 = traditional quizzes, 1 = gamified quizzes. Quizzes Completed and the Course Grade were z-standardized for these calculations. The 95% confidence interval for the effects is obtained by the bias-corrected bootstrap with 10,000 resamples. $N_{TestA} = 454$, $N_{TestB} = 444$, $N_{TestC} = 437$.
*p < .05, **p < .01.

Table 3
Results for the conditional direct and indirect effects of the condition on the test scores.

Model	Course Grade	Effect	*SE*	95% Confidence Interval
Conditional Direct Effects				
Quiz Format → Score Test A	Low	-.11	.12	[-.36, .13]
Quiz Format → Score Test A	Medium	.40	.09	[.21, .58]
Quiz Format → Score Test A	High	.87	.13	[.62, 1.12]
Quiz Format → Score Test B	Low	-.01	.13	[-.27, .24]
Quiz Format → Score Test B	Medium	.28	.10	[.09, .47]
Quiz Format → Score Test B	High	.54	.13	[.28, .79]
Quiz Format → Score Test C	Low	-.12	.13	[-.38, .14]
Quiz Format → Score Test C	Medium	.28	.10	[.09, .47]
Quiz Format → Score Test C	High	.67	.14	[.39, .96]
Conditional Indirect Effects				
Quiz Format → Quizzes Completed → Score Test A	Low	-.06	.03	[-.15, −.01]
Quiz Format → Quizzes Completed → Score Test A	Medium	-.05	.02	[-.11, −.02]
Quiz Format → Quizzes Completed → Score Test A	High	.04	.03	[-.11, .01]
Quiz Format → Quizzes Completed → Score Test B	Low	-.05	.04	[-.12, .01]
Quiz Format → Quizzes Completed → Score Test B	Medium	-.01	.02	[-.06, .03]
Quiz Format → Quizzes Completed → Score Test B	High	.03	.04	[-.04, .13]
Quiz Format → Quizzes Completed → Score Test C	Low	.01	.01	[-.03, .09]
Quiz Format → Quizzes Completed → Score Test C	Medium	.02	.02	[-.02, .06]
Quiz Format → Quizzes Completed → Score Test C	High	.03	.04	[-.02, .13]

Note. Coding of the variable Quiz Format: 1 = traditional quizzes, 1 = gamified quizzes. Values of the moderator levels for the moderator Students' Course Grade: Low = minus one SD from the mean, Medium = mean, High = plus one SD from the mean. Quizzes Completed and Course Grade were z-standardized for these calculations. The 95% confidence interval for the effects is obtained by the bias-corrected bootstrap with 10,000 resamples. *SE* = standard error of the effect sizes. $N_{TestA} = 454$, $N_{TestB} = 444$, $N_{TestC} = 437$.

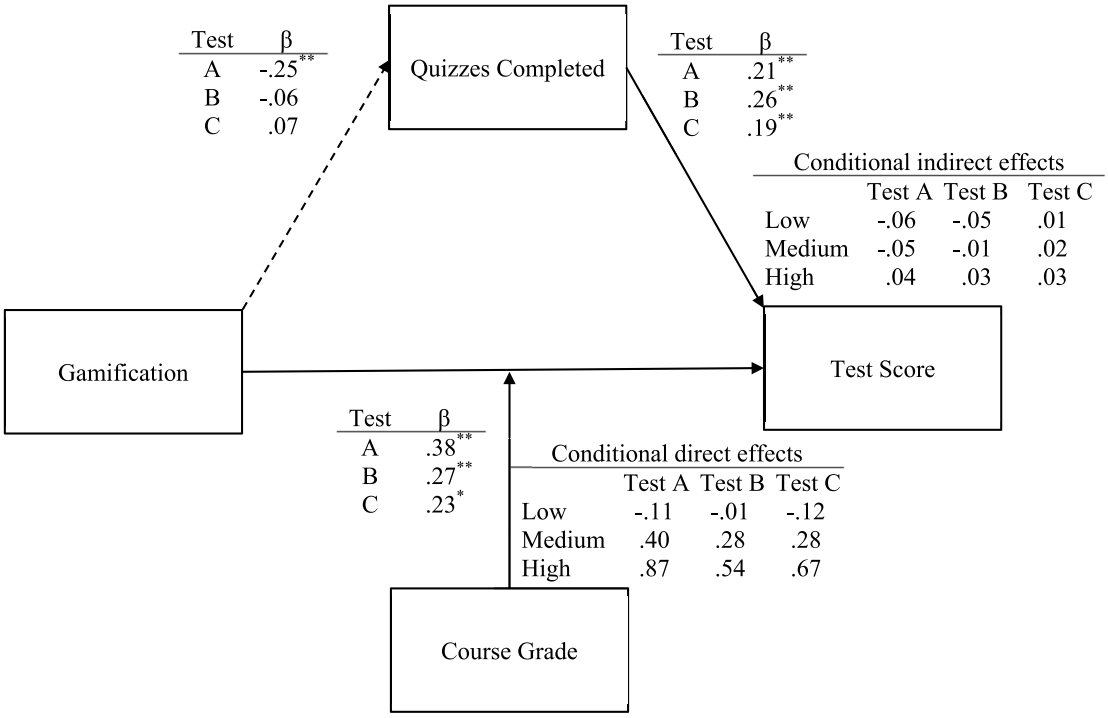

Fig. 3. Results of the PROCESS moderated-mediation analysis. The dotted line indicates that there was only a significant effect of the Quiz Format on Quizzes Completed for Test A. Low = Low Course Grade, Medium = Medium Course Grade, High = High Course Grade. $N_{TestA} = 454$, $N_{TestB} = 444$, $N_{TestC} = 437$. $^{*}p < .05$, $^{**}p < .01$.

for higher and lower achieving students). Similar results were found for the complete models testing the effect of the number of Quizzes Completed on the Test Scores, as there were no interactions of the number of Quizzes Completed and the Course Grade. This implies that Course Grade did not moderate the effect of number of Quizzes Completed on the test scores (i.e., number of Quizzes Completed did not have a different effect on Test Scores for higher and lower achieving students).

4. Discussion

In this study we applied the theory of gamified learning to structure our approach in understanding gamified quizzes in an educational setting. In this comparison, students completed gamified or traditional online quizzes in preparation for class tests. The goals were to replicate the testing effect in an online environment, test if gamification can enhance the testing effect, and to determine if there are student characteristics that affect the effects of gamification. The results show that the testing effect has a robust effect on learning through online quizzes, that gamification did not enhance the testing effect, and that student abilities can impact outcomes of gamification.

First, we were able to replicate the testing effect supporting that our sample does hold to previous research findings (Rowland, 2014). The more quizzes a student took regardless of the Quiz Format, the better they did on the individual tests throughout the semester. This finding supports the positive implications of the testing effect on student learning (see e.g. Jensen et al., 2014; Johnson & Mayer, 2009) and provides further support that the testing effect works in online settings.

Furthermore, we found that there was a direct effect of gamification on the Test Scores for the initial test (i.e., Test A) but not for subsequent tests (i.e., Test B, and C). We explicitly examined the effects of gamification for learning on the three tests over the course of the semesters (instead of combining all test scores into one overall score), as this made it possible to examine if the beneficial effects of gamification might wear off over time. This way, the current findings respond to calls for research investigating possible novelty effects of gamification through longitudinal designs providing additional empirical evidence for the previously stated concerns raised by other researchers that the primary benefit of gamification may be a novelty effect that wears off over time (Hamari et al., 2014). One possible explanation for this effect is that perceived enjoyment and usefulness of a gamified activity diminishes (see also Hamari et al., 2014), which might affect benefits from a gamified activity. Supporting this explanation, Hamari et al. (2014) examined 24 studies and found a general positive effect of gamification on learner behavior and psychological outcomes (e.g., motivation). However, they stressed that some of the positive effects may have been due to a novelty effect rather than the long-term success of gamification and affirmed the possible presence of the novelty effect in one of their follow-up studies (Koivisto & Hamari, 2014). They found that perceived enjoyment and usefulness of gamification declined with use and suggested that a novelty effect caused the initial benefits of gamification. An important implication following this discussion and the findings of the current study is

that gamification might be a short-time impulse rather than an effective tool for sustainable changes in pedagogical environments (Koivisto & Hamari, 2014). Moreover, the current results contribute to the discussion on novelty effects as they indicate that there are circumstances where the novelty effect already wears off after a short time-span and show that a potential novelty effect might especially affect proposition 3 of the theory of gamified learning (i.e., "game characteristics influence changes in behavior/attitudes"; Landers, 2014, p. 761).

In contrast to our theoretical assumptions, gamification did not positively impact the number of Quizzes Completed. Following Landers (2014) and Landers et al. (2017) (who engaged students through leaderboards) this could have been one way gamification improves learning; through more student engagement which might have benefitted the testing effect. In the theoretical background, we showed that our version of gamification includes game elements based on the same game characteristics like the leaderboards used by Landers and Landers (2014). Specifically, the game elements of the current study also reflected the game characteristics assessment and rules/goals. Accordingly, we assumed that our game elements should lead to students completing more quizzes reflecting increases in student engagement. However, additionally to assessment and rules/goals, leaderboards also include the game characteristic conflict/challenge. In the case of the study of Landers and Landers (2014), this was competition between students to be on top of the leaderboard. It could be possible that it was this game characteristic that engaged their students to take part in learning activities and that would also have been one way to engage the students in the current study to complete more quizzes. In other words, conflict/challenge could be the determining aspect that led to student engagement within the study of Landers and Landers (2014). Yet, this assumption calls for research determining if leaderboards can actually be an effective way of strengthening the testing effect by motivating students to take more tests to outperform other students on the tests.

In contrast to our theoretical assumptions, findings for the Quizzes Completed prior to Test A indicate that students in the gamified quiz condition took part in fewer quizzes than students who experienced traditional quizzes. Yet, they outperformed the traditional quiz group in Test A. This could mean that instead of engaging students to complete more quizzes (Werbach & Hunter, 2012), the current version of gamification might have had a different effect on students. Possibly, the game elements created a more rewarding experience of completing the quizzes (cf. Kwon, Halavais, & Havener, 2015). Specifically, the encouraging messages might have felt rewarding (Zichermann & Cunningham, 2011), the progress bar could have shown students that they successfully advance their knowledge (Dicheva et al., 2015), and not needing the hint button any more might have also constituted a rewarding experience. This kind of might have also initiated effects other than what we initially assumed in the theoretical background. For instance, the encouraging messages may have provided too good of a positive feedback, which consequently may have given students the confidence that they were ready for the exam and prevented them from returning to the quiz. Similarly, the progress bar may have also generated a sense of completion where students felt like they were finished with the quizzes and didn't need to complete them again. It is possible that when students saw a full progress bar they might perceive that they mastered this quiz (cf. Dicheva et al., 2015). Lastly, if students opt out of the wager option, they may also interpret this as a message of preparedness for the tests.

We interpret this effect in a way that students who received these additional features on their progress and performance (i.e., more explicit than in the traditional quiz condition) may have perceived that they achieved everything needed to prepare for the test using the quizzes, thus leading them to complete fewer quizzes (at least for Test A). The game elements potentially increased student's level of self-efficacy regarding the learning content and this higher self-efficacy might have positively affected test performance (cf., Tay, Ang, & Van Dyne, 2006). Supporting this assumption, Gist and Mitchell (1992) present an overview of determinants of self-efficacy showing that feedback and mastery experience can affect self-efficacy. In the case of the current study, encouraging messages as well as the progress bar and becoming independent from the hints might all account for these feedback and mastery processes to influence self-efficacy (cf. Dicheva et al., 2015). However, we advise caution regarding this interpretation of the results as this result was only found for Test A, and as there is strong need for future studies to support the aforementioned assumption.

Furthermore, the results of this study contribute to growing research on the influence of individual differences on the effects of gamification (e.g., Landers & Armstrong, 2017). We found that the available demographic variables (i.e., gender and year in school) did not affect the interpretation of our findings when used as covariates. However, lower achieving students benefited less from the potential positive effects of gamification expanding research on the moderating effects of individual differences on gamification (e.g., Koivisto & Hamari, 2014; Landers & Armstrong, 2017). Additionally, the results of the moderated-mediation imply that there might be long-lasting effects of gamification for higher achieving students which is in contrast to the previously discussed novelty effect. Specifically, we found that there was a direct positive effect of gamification on all the tests for students with medium and high course grades. In combination with the finding that gamification did not affect the number of Quizzes Completed, the direct positive effect of gamification on learning for higher achieving students might indicate that these students benefited from gamification in a way that was not captured in this study. As already discussed before, students' self-efficacy might be the variable that accounts for the findings of the current study – more precisely, self-efficacy resulting from a mastery experience for the quizzes (cf. Dicheva et al., 2015; Gist & Mitchell, 1992). It is imaginable that the gamified version of the quizzes especially improved self-efficacy only for students with already higher self-efficacy because of previously experiencing success in their studies (cf., Bandura, 2006).

Another possible explanation for the differential effects based on students' abilities might come from research regarding seductive details (Rey, 2012). This research has shown that adding illustrative but irrelevant information to a text or a presentation (i.e., seductive details) can distract learners from the actual comprehension of the content (Garner, Brown, Sanders, & Menke, 1992). This seems to be especially true for people with lower working-memory capacities (Sanchez & Wiley, 2006). We used students' course grade as a measure for students' capabilities to study at a university. Hence, it might also be possible that course grade is correlated with students' cognitive abilities such as working-memory capacity (Richardson et al., 2012). This could mean that some game features may be more of a distraction than a benefit to lower achieving students which would correspond to research on seductive details. However, investigating if certain approaches to gamification really lead to effects similar to detrimental effects found for

seductive details would need to be thoroughly investigated in future studies.

A potential consequence of the aforementioned effects and the moderating role of students' abilities on the benefits of gamification might be a Matthew effect (Penno, Wilkinson, & Moore, 2002) in a way that higher achieving students benefit more from gamification whereas lower achieving students are left behind. This is especially problematic in cases when instructors consider applying gamification to specifically enhance motivation in lower achieving students as it could increase the gap between higher and lower achieving students. It is important to note that this study is not the first one to point at the potential negative effects of gamification. For instance, Hanus and Fox (2015) found that gamification through leaderboards and badges led to less motivation and lower course grades. In our case, gamification through points did not lead to positive effects for lower achieving students. Therefore, this finding calls for further research regarding antecedents and conditions for successful gamification (i.e., the reasons for why lower achieving students did not benefit from gamification).

Interpersonal differences seem to be especially important when trying to predict if and for whom gamification has positive effects (see also Koivisto & Hamari, 2014). For instance, a tentative interpretation of the findings of the current study would be that certain ways to gamify learning might work better for people with higher cognitive abilities. There are many other interpersonal differences that could possibly affect the effects of gamification. Since gamification elements such as leaderboards work through competition, it might also be interesting to investigate the effects of people's competitiveness (see e.g., Smither & Houston, 1992) on gamification and the list of variables potentially affecting gamification can be extended easily (e.g., Big Five Personality factors, locus of control).

4.1. Limitations

There are at least three limitations that need to be addressed in the current study. First, due to having real-world field data, this study is based on a quasi-experimental design. This means that students could not be randomly assigned to the different Quiz Formats due to the nature of the data collection (i.e., assessing two subsequent semesters). Thus, in this particular comparison, findings should be interpreted cautiously as there might be systematic differences between the courses that were not canceled out by randomization (e.g., students who take an introduction to psychology class in the fall rather than the spring may have systematic differences from one another that could not be controlled for). However, many elements were consistent between the two semesters (e.g., the same instructor, same pool of quizzes, same tests), and including Gender as well as year of study as covariates did not affect the interpretation of the results.

Second, the gamification approach used three game elements. Therefore, it is difficult to tell which of these components had which effects. It is likely that displaying overall scores with an encouraging message as the only gamified feature would have different effects than presenting a progress bar with current and potential points as the only gamified feature. Future studies should be aware of this potential issue with game elements that was also highlighted by Landers et al. (2018) who stated that it is necessary to also investigate the effects of single gamification elements in order to understand their effects.

Third, we aimed at revealing a potential psychological effect through which gamification should enhance the testing effect in a way that gamification should enhance motivation to complete preparational quizzes. However, the results show that gamification worked (for Test A and for higher achieving students) through a different mechanism. Therefore, we were not able to open the black box through which the current version of gamification affected learning. Future studies could consider gathering qualitative data from participants which would help to reveal students' thoughts on gamification and provides insights into their learning experience through gamification. Even though the current study was not able to open the black-box on the effects of gamification on learning, its results contribute to research and practice highlighting the need for more research on the attendance to the positive effects of gamification (e.g., individual differences). Therefore, this study repeats and reinforces calls (e.g., Landers et al., 2018) for extensive research on revealing the psychological underpinnings of the effects of gamification.

4.2. Practical implications

For instructors there might be a variety of take-aways from the current study. First, the current form of gamification did initially benefit student learning. Since the game elements of the current study can be easily included in existing instructional content, instructors might consider similar forms of gamification for enhancing short-term assignments and activities to improve learning outcomes. Second, it seems not to be effective to use the same game elements to enhance learning permanently or for long-term assignments. Similar to previous research (e.g., Koivisto & Hamari, 2014), there seems to be evidence that gamification is not sustainably improving student learning. Given this, game features may lose their influence already after a short period of time. Thus, the game features possibly would need to be altered to once again be novel to continue to show improvements over traditional methods. Lastly, game features should be introduced with care, considering that there seem to be contexts where higher achieving students benefit more than lower achieving ones. Although this is not reason to rule out gamification, it does call for special considerations and designs meeting the needs of this population. Possibly, some game features may be more of a distraction than a benefit to lower achieving students which would correspond to research on seductive details (cf., Sanchez & Wiley, 2006). Thus, our final advice to instructors is to approach applications of gamification with intention, understanding that these features can be of use but are best designed for the particular assignment and group of students who will be completing them.

4.3. Future research

The current study reinforces previous calls for research in the area of gamification (e.g., Landers et al., 2018). Future work could

develop ideas to enhance the sustainability of gamification effects. One fruitful avenue could be to investigate the effect of constant changes in game elements. For instance, in the beginning of a course, students could receive points for assignments. After some time, it might be possible to add a leaderboard in which students can compete with each other. After more time, the leaderboard could include features like badges to motivate students to fulfill certain tasks. This would closely follow the model of common games (e.g., trading card games; MMORPGs) which constantly add new game features and change game rules and elements in order to keep gamers motivated (see e.g., Hodge et al., 2018).

Another consideration for future research comes from findings in the current study which point towards interpersonal differences as factors influencing the effects of gamification. It is equally likely that situational factors influence gamification. For instance, the same gamification approach might affect learning in a pedagogical setting positively whereas it does not affect learning in a workplace setting. One reason for such a finding could be that gamification seems appropriate in one setting (e.g., motivating students to do their homework) but not in another (e.g., motivating employees to take part in training).

5. Conclusion

The current study replicated the testing effect in an online learning setting, implied that gamification can have short-term positive effects, and showed that interpersonal differences can influence the positive effects of gamification. While our findings imply that the effects of certain gamification approaches seem to be short-lived and may only be beneficial to higher performing individuals, our findings also strongly support the call for more scrutiny regarding antecedents and consequences of, as well as psychological processes behind gamification (Landers et al., 2018) as there are still many things hidden inside of black boxes ready to be examined by future research.

References

Armstrong, M. B., Landers, R. N., & Collmus, A. B. (2016). Gamifying recruitment, selection, training, and performance management: Game-thinking in human resource management. In H. Gangadharbatla, & D. Z. Davis's (Eds.). *Emerging research and trends in gamification* (pp. 140–165). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-4666-8651-9.

Baird, L. L. (1985). Do grades and tests predict adult accomplishment? *Research in Higher Education, 23*(1), 3–85. https://doi.org/10.1007/BF00974070.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-Efficacy Beliefs of Adolescents, 5*(1), 307–337.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *The Journal of Educational Research, 85*(2), 89–99. https://doi.org/10.1080/00220671.1991.10702818.

Barkley, E. F., Cross, K. P., & Major, C. H. (2014). *Collaborative learning techniques: A handbook for college faculty.* San Francisco, CA: John Wiley & Sons.

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a taxonomy linking game attributes to learning: An empirical study. *Simulation & Gaming, 43*(6), 729–760. https://doi.org/10.1177/1046878112439444.

Bonk, C. J., & Graham, C. R. (2012). *The handbook of blended learning: Global perspectives, local designs.* San Francisco, CA: John Wiley & Sons.

Boticki, I., Baksa, J., Seow, P., & Looi, C. K. (2015). Usage of a mobile social learning platform with virtual badges in a primary school. *Computers & Education, 86*, 120–136. https://doi.org/10.1016/j.compedu.2015.02.015.

Brull, S., & Finlayson, S. (2016). Importance of gamification in increasing learning. *The Journal of Continuing Education in Nursing, 47*(8), 372–375. https://doi.org/10.3928/00220124-20160715-09.

Buabeng-Andoh, C. (2012). Factors influencing teachers' adoption and integration of information and communication technology into teaching: A review of the literature. *International Journal of Education and Development Using Information and Communication Technology, 8*(1), 136–155.

Canhoto, A. I., & Murphy, J. (2016). Learning from simulation design to develop better experiential learning initiatives an integrative approach. *Journal of Marketing Education, 38*(2), 98–106. https://doi.org/10.1177/0273475316643746.

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science, 21*(5), 279–283. https://doi.org/10.1177/0963721412452728.

Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition, 34*(2), 268–276. https://doi.org/10.3758/BF03193405.

Chang, W., & Taylor, S. A. (2016). The effectiveness of customer participation in new product development: A meta-analysis. *Journal of Marketing, 80*(1), 47–64. https://doi.org/10.1509/jm.14.0057.

Chang, J. W., & Wei, H. Y. (2016). Exploring engaging gamification mechanics in massive online open courses. *Journal of Educational Technology & Society, 19*(2), 177–203.

Coccoli, M., Iacono, S., & Vercelli, G. (2015). Applying gamification techniques to enhance effectiveness of video-lessons. *Journal of e-Learning and Knowledge Society, 11*(3), 73–84.

Cornelissen, F., Neerincx, M. A., Smets, N., Breebaart, L., Dujardin, P., & Wolff, M. (2013). Gamification for astronaut training. In M. Schmidhuber, C. Cruzen, & J. Kehr's (Eds.). *Progress in astronautics and aeronautics space operations: Experience, mission systems, and advanced concepts* (pp. 91–110). American Institute of Aeronautics and Astronautics, Inc. https://doi.org/10.2514/5.9781624102080.0091.0110.

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011, May). Gamification: Toward a definition. *CHI 2011 gamification workshop proceedings: Vol. 12* (Vancouver, BC, Canada).

Di Bitonto, P., Corriero, N., Pesare, E., Rossano, V., & Roselli, T. (2014, June). Training and learning in e-health using the gamification approach: The trainer interaction. In C. Stephanidis, & M. Antona (Eds.). *Universal access in human-computer interaction (UAHCI) aging and assistive environments. (8515)* (pp. 228–237). Cham: Springer. https://doi.org/10.1007/978-3-319-07446-7_22.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society, 18*(3), 75–88.

Dicheva, D., Irwin, K., Dichev, C., & Talasila, S. (2014). A course gamification platform supporting student motivation and engagement. *Presented at the IEEE international Conference on Web and open Access to learning. Dubai, UAE*https://doi.org/10.1109/icwoal.2014.7009214.

Domingo, M. G., & Garganté, A. B. (2016). Exploring the use of educational technology in primary education: Teachers' perception of mobile technology learning impacts and applications' use in the classroom. *Computers in Human Behavior, 56*, 21–28. https://doi.org/10.1016/j.chb.2015.11.023.

Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research, 12*(7), 2121–2159.

Farcas, S., Szamosközi, I., & Takacs, A. (2016). The effects of working memory trainings with game elements for children with ADHD. A meta-analytic review. *Transylvanian Journal of Psychology, 1*(2), 21–44.

Findley, D. A., Keever, G. J., Chappelka, A. H., Eakes, D. J., & Gilliam, C. H. (1996). Sensitivity of red maple cultivars to acute and chronic exposures of ozone. *Journal of Arboriculture, 22*, 264–269.

Garner, R., Brown, R., Sanders, S., & Menke, D. J. (1992). "Seductive details" and learning from text. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.). *The role of interest in learning and development* (pp. 239–254). Hillsdale, NJ: Erlbaum.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441–467. https://doi.org/10.1177/1046878102238607.

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education, 19*(3), 133–148. https://doi.org/10.1207/s15389286ajde1903_2.

Gebhard, P., Schneeberger, T., Andre, E., Baur, T., Damian, I., Mehlmann, G., et al. (2018). Serious games for training social skills in job interviews. *IEEE Transactions on Games, 1*. https://doi.org/10.1109/tg.2018.2808525.

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183–211. https://doi.org/10.5465/amr.1992.4279530.

Goh, T.-T., Seet, B.-C., & Chen, N.-S. (2012). The impact of persuasive SMS on students' self-regulated learning. *British Journal of Educational Technology, 43*(4), 624–640. https://doi.org/10.1111/j.1467-8535.2011.01236.x.

Gray, H. M., LaPlante, D. A., & Shaffer, H. J. (2012). Behavioral characteristics of Internet gamblers who trigger corporate responsible gambling interventions. *Psychology of Addictive Behaviors, 26*(3), 527–535. https://doi.org/10.1037/a0028545.

Green, H., & Hannon, C. (2007). *Young people are spending their time in a space which adults find difficult to supervise or understand.* (London, England).

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?–A literature review of empirical studies on gamification. *Proceedings of the 47th Hawaii international Conference on system sciences (HICSS), Hawaii, USA*https://doi.org/10.1109/HICSS.2014.377.

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior, 54*, 170–179. https://doi.org/10.1016/j.chb.2015.07.045.

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 & Education, 80*, 152–161. https://doi.org/10.1016/j.compedu.2014.08.019.

Harvey, W. L. (1970). *A study of the cognitive and affective outcomes of a collegiate science learning game.* (Doctoral dissertation). Available from Dissertation Abstracts International, (ED050552).

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research, 50*(1), 1–22. https://doi.org/10.1080/00273171.2014.962683.

Herrington, J., Oliver, R., & Reeves, T. C. (2003). Patterns of engagement in authentic online learning environments. *Australasian Journal of Educational Technology, 19*(1), https://doi.org/10.14742/ajet.1701.

Hodge, S. E., McAlaney, J., Gatzidis, C., Anderson, E. F., Melacca, D., & Taylor, J. (2018). Applying Psychological Theory to in-game moral behaviors through the development of a purpose-made game. In N. D. Bowman (Ed.). *Video games: A medium that demands our attention* (pp. 108–125). New York: Routledge.

Irving, K., Pape, S., Owens, D., Abrahamson, L., Silver, D., & Sanalan, V. (2016). Classroom connectivity and algebra 1 achievement: A three-year longitudinal study. *Journal of Computers in Mathematics and Science Teaching, 35*(2), 131–151.

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review, 26*(2), 307–329. https://doi.org/10.1007/s10648-013-9248-9.

Johnson, C. I., & Mayer, R. E. (2009). A testing effect with multimedia learning. *Journal of Educational Psychology, 101*(3), 621–629. https://doi.org/10.1037/a0015183.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341–350. https://doi.org/10.1037/0003-066X.39.4.341.

Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. *The Internet and Higher Education, 22*, 37–50. https://doi.org/10.1016/j.iheduc.2014.04.003.

Kling, N., McCorkle, D., Miller, C., & Reardon, J. (2005). The impact of testing frequency on student performance in a marketing course. *The Journal of Education for Business, 81*(2), 67–72. https://doi.org/10.3200/JOEB.81.2.67-72.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior, 35*, 179–188. https://doi.org/10.1016/j.chb.2014.03.007.

Koka, A., & Hein, V. (2003). Perceptions of teacher's feedback and learning environment as predictors of intrinsic motivation in physical education. *Psychology of Sport and Exercise, 4*(4), 333–346. https://doi.org/10.1016/s1469-0292(02)00012-2.

Kwon, K. H., Halavais, A., & Havener, S. (2015). Tweeting badges: User motivations for displaying achievement in publicly networked environments. *Cyberpsychology, Behavior, and Social Networking, 18*(2), 93–100. https://doi.org/10.1089/cyber.2014.0438.

Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming, 45*(6), 752–768. https://doi.org/10.1177/1046878114563660.

Landers, R. N., & Armstrong, M. B. (2017). Enhancing instructional outcomes with gamification: An empirical test of the Technology-Enhanced Training Effectiveness Model. *Computers in Human Behavior, 71*, 499–507. https://doi.org/10.1016/j.chb.2015.07.031.

Landers, R. N., Auer, E. M., Collmus, A. B., & Armstrong, M. B. (2018). Gamification science, its history and future: Definitions and a research agenda. *Simulation & Gaming, 49*(3), 315–337. https://doi.org/10.1177/1046878118774385.

Landers, R. N., Bauer, K. N., & Callan, R. C. (2017). 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., & Callan, R. C. (2011). Casual social games as serious games: The psychology of gamification in undergraduate education and employee training. In M. Ma, A. Oikonomou, & L. C. Jain (Eds.). *Serious games and edutainment applications* (pp. 399–423). London: Springer London. https://doi.org/10.1007/978-1-4471-2161-9_20.

Landers, R. N., & Landers, A. K. (2014). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming, 45*(6), 769–785. https://doi.org/10.1177/1046878114563662.

Langer, M., König, C. J., Gebhard, P., & André, E. (2016). Dear computer, teach me manners: Testing virtual employment interview training. *International Journal of Selection and Assessment, 24*(4), 312–323. https://doi.org/10.1111/ijsa.12150.

Larsen, D. P. (2013). When I say test-enhanced learning. *Medical Education, 47*(10), https://doi.org/10.1111/medu.12238 961-961.

Lister, C., West, J. H., Cannon, B., Sax, T., & Brodegard, D. (2014). Just a fad? Gamification in health and fitness apps. *JMIR Serious Games, 2*(2), https://doi.org/10.2196/games.3413.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., et al. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology, 34*(1), 51–57. https://doi.org/10.1016/j.cedpsych.2008.04.002.

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. L., III (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology, 103*(2), 399–414. https://doi.org/10.1037/a0021782.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007a). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*(4–5), 494–513. https://doi.org/10.1080/09541440701326154.

McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007b). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14*(2), 200–206. https://doi.org/10.3758/BF03194052.

Means, B. (2010). Technology and education change: Focus on student learning. *Journal of Research on Technology in Education, 42*(3), 285–307. https://doi.org/10.1080/15391523.2010.10782552.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*Report No. ED-04-CO-0040. Washington, DC: US Department of Education.

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.

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior, 23*(3), 1126–1144. https://doi.org/10.1016/j.chb.2006.10.006.

Oprescu, F., Jones, C., & Katsikitis, M. (2014). I play at work—ten principles for transforming work processes through gamification. *Frontiers in Psychology, 5*(14), 1–5. https://doi.org/10.3389/fpsyg.2014.00014.
Orvis, K. A., Horn, D. B., & Belanich, J. (2009). An examination of the role individual differences play in videogame–based training. *Military Psychology, 21*(4), 461–481. https://doi.org/10.1080/08995600903206412.
Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94*(1), 23.
Reigeluth, C. M. (1983). *Instructional design theories and models: An overview of their current status.* Hillsdale, NJ: Erlbaumhttps://doi.org/10.4324/9780203824283.
Rey, G. D. (2012). A review of research and a met-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216–237. https://doi.org/10.1016/j.edurev.2012.05.003.
Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838.
Richardson, J., & Swan, K. (2003). *Examining social presence in online courses in relation to students' perceived learning and satisfaction.* Doctoral dissertation, State University of New York at Albanyhttp://hdl.handle.net/2142/18713.
Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal, 53*(3), 617–635. https://doi.org/10.5465/amj.2010.51468988.
Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. https://doi.org/10.1037/a0037559.
Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Ego-involved persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion, 15*(3), 185–205. https://doi.org/10.1007/BF00995170.
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.
Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition, 34*(2), 344–355. https://doi.org/10.3758/bf03193412.
Sandholtz, J. H. (1997). *Teaching with technology: Creating student-centered classrooms.* New York, NY: Teachers College Press.
Singh, S. P. (2012). Gamification: A strategic tool for organizational effectiveness. *International Journal of Management, 1*(1), 108–113. https://doi.org/10.15410/aijm/2012/v1i1/50480.
Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology, 59*(3), 623–664. https://doi.org/10.1111/j.1744-6570.2006.00049.x.
Smither, R. D., & Houston, J. M. (1992). The nature of competitiveness: The development and validation of the competitiveness index. *Educational and Psychological Measurement, 52*(2), 407–418. https://doi.org/10.1177/0013164492052002016.
Su, C. H., & Cheng, C. H. (2015). A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning, 31*(3), 268–286. https://doi.org/10.1111/jcal.12088.
Tay, C., Ang, S., & Van Dyne, L. (2006). Personality, biographical characteristics, and job interview success: A longitudinal study of the mediating effects of interviewing self-efficacy and the moderating effects of internal locus of causality. *Journal of Applied Psychology, 91*(2), 446–454. https://doi.org/10.1037/0021-9010.91.2.446.
Vojdanoska, M., Cranney, J., & Newell, B. R. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology, 24*(8), 1183–1195. https://doi.org/10.1002/acp.1630.
Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM, 51*(8), 58–67. https://doi.org/10.1145/1378704.1378719.
Wells, C. (2016). To game or not to game: An investigation of the impact of survey visualisation and gamification. *International Journal of Market Research, 58*(2), 325–331. https://doi.org/10.2501/IJMR-2016-017.
Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business.* Philadelphia, PA: Wharton Digital Press.
Wheeler, M. A., & Roediger, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science, 3*(4), 240–246. https://doi.org/10.1111/j.1467-9280.1992.tb00036.x.
Wing, E. A., Marsh, E. J., & Cabeza, R. (2013). Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect. *Neuropsychologia, 51*(12), 2360–2370. https://doi.org/10.1016/j.neuropsychologia.2013.04.004.
Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps.* Sebastopol, CA: O'Reilly Media, Inc.
Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663–676. https://doi.org/10.2307/1163261.