Article

Simulation & Gaming
1–17

DOI: 10.1177/1046878114563662
sag.sagepub.com

An Empirical Test of the Theory of Gamified Learning: The Effect of Leaderboards on Time-on-Task and Academic Performance

Richard N. Landers[1] and Amy K. Landers[1]

Abstract

Background. The theory of gamified learning provides a theoretical framework to test the **impact** of **gamification** efforts upon learner **behaviors** and **attitudes**, as well as the effect of these behavioral and attitudinal changes on learning. It does so by providing **mediating** and **moderating** processes **linking** specific **game elements** to **learning outcomes**.

Aim. This article links specific game elements common to **leaderboards** (conflict/challenge, rules/goals, and assessment) with a focal **learner behavior**, time-on-task, by exploring educational research on **competition** and psychological research on **goal-setting** theory.

Method. The mediating process of the theory of gamified learning is tested **experimentally** by assigning learners completing an online wiki-based project to a **gamified version with a leaderboard** or to a **control version without a leaderboard**. Leaderboard achievement was not tied to course grades.

Results. Random assignment to leaderboards supported a **causal effect**. Students with **leaderboards** interacted with their project 29.61 **more** times, on average, than those in a control condition. **Bootstrapping** was used to support the mediation of the effect of gamification on academic achievement by this amount of time.

Conclusion. The **mediating process** of the theory of gamified instruction is **supported**. **Leaderboards** can be used to **improve** course **performance** under certain **circumstances**.

[1]Old Dominion University, USA

Corresponding Author:
Richard N. Landers, Old Dominion University, 250 Mills Godwin Building, Norfolk, VA 23529, USA.
Email: rnlanders@odu.edu

Keywords
attitudes, behavior, game attribute taxonomy, game attributes, game element taxonomy, game elements, gamification, gamified learning, learning, learning outcomes, mediation, model, moderation, psychology, serious games, simulation/gaming, taxonomy, theory, training

Landers (in press) introduced the theory of gamified learning to explain the causal paths by which gamification interventions can affect outcomes for learners across a wide variety of contexts. This theory describes two specific processes by which gamification can be used to affect learning outcomes: a mediating process and a moderating process. Incomplete consideration of the psychological processes involved can result in gamification appearing to fail to achieve intended results (Hamari, Koivisto, & Sarsa, 2014). However, failures can in fact occur at multiple points, and the fault is only sometimes due to design. This further implies that researchers working to build a scientific literature of gamification must identify the target process carefully and only then decide upon an appropriate research design and rigorous evaluation strategy of that process.

First, gamification was theorized to affect learning via a mediating process wherein gamification alters a psychological characteristic that itself affects an outcome of interest (see Figure 1; D $\rightarrow$ C $\rightarrow$ B). For example, the game attribute *conflict/challenge* (see Bedwell, Pavlas, Heyne, Lazzara, & Salas, 2012) might be used in a learning activity to create the experience of time pressure (in order to better represent the real-world task it is modeled upon), and the experience of time pressure might itself improve transfer of gained skills outside of the classroom. In this example, the use of conflict/challenge must successfully create the experience of time pressure, and the experience of time pressure must lead to improved transfer for gamification to be considered successful.

Second, gamification was theorized to affect learning via a moderating process wherein gamification alters a psychological characteristic that strengthens the relationship between the quality of instructional content and learning outcomes (see Figure 1; D $\rightarrow$ C which moderates A $\rightarrow$ B). For example, the game attribute *game fiction* might be used to increase learner engagement, which is intended to increase learner motivation to maintain their attention on provided instructional content. In this example, game fiction must successfully increase learner engagement, and learner engagement must strengthen the relationship between instructional content and outcomes.

Landers (in press) calls for rigorous empirical tests of both theoretical processes using both correlational and experimental approaches, and we choose to focus in this article on the mediating process for three reasons. First, it provides the most direct path to improve learning outcomes in the context of education and training. Even in the absence of a moderating effect, the presence of a mediating effect would suggest gamification could causally improve learning. Second, as a consequence of its more indirect path to affect learning, the magnitude of the moderating effect is likely smaller

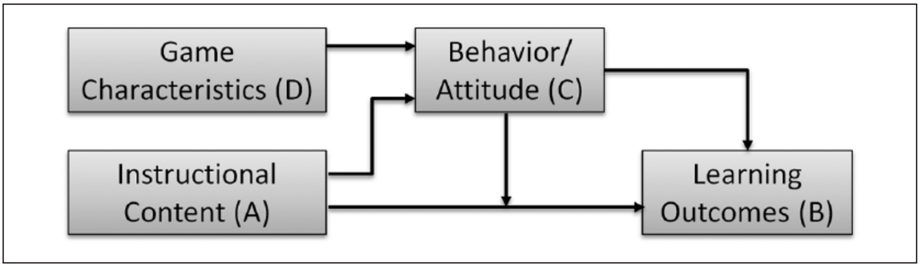

Figure 1. Theory of gamified learning.
Note. D → C → B is the mediating process described by this theory. The influence of C on A → B is the moderating process described by this theory. Directional arrows indicate theorized path of causality.

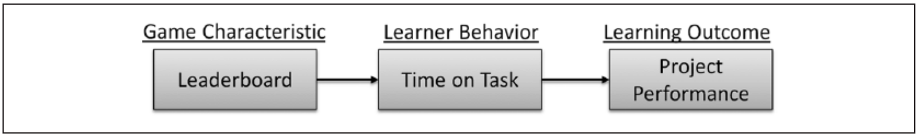

Figure 2. Mediating process of the theory of gamified learning tested in this study.

than the mediating effect. Targeting the mediating effect allows us to explain more variance in learning outcomes first, leaving smaller amounts of variance to be explained in subsequent research. Third, we wished to use an experimental design in a field sample to maximize both internal and ecological validity (Cooke, as cited in Wolfe & Crookall, 1998), and manipulating instructional quality (e.g., randomly assigning learners to poor-quality and high-quality instruction) as required by experimental tests of the moderating process is not ethical.

Given this, we identified a course in which a specific behavioral outcome potentially alterable by gamification might be used to benefit student learning outcomes with the purpose of empirically testing the mediating relationship between specific game elements common to leaderboards, a target behavior, and academic performance. In previous semesters of this course, one of the most significant challenges faced by the instructor of the course we identified was time spent on the course's term project. A full semester was provided so that students would work on their project frequently and throughout the semester as a learning experience, and this interaction with the material over time was a primary aspect of the project theorized to improve learning. To the dismay of the instructor and reflecting a general tendency among undergraduates toward procrastination (Pychyl, Lee, Thibodeau, & Blunt, 2000), most students would wait until the end of the semester to work on the project, usually a day or two before it was due. It is in this context that we identified leaderboards as (a) a potential solution to the instructor's problem, (b) a valid test of the mediating process of the theory of gamified learning (see Figure 2), and (c) a relatively common educational gamification intervention (Landers & Callan, 2011). Specifically, we sought to demonstrate that leaderboards could be used to increase student achievement by encouraging them to spend more time on the project than they would have without leaderboards.

Time-on-Task as a Cause of Learning

A review of prior research in both education and workplace learning reveals that increased time-on-task increases learning across all learning contexts (Brown, 2001). From a century of research, it is now accepted that learners who practice and engage with a task more often produce greater knowledge and skill than those who do not practice (Ericsson, Krampe, & Tesch-Romer, 1993; Goldstein, 1993). The underlying processes involved have been explored in research on resource allocation models of human attention; people who keep their attention focused upon a task learn more than those who do not (Kanfer & Ackerman, 1989). Many serious games researchers cite increasing student time-on-task as one of the primary motivations for incorporating games in the classroom (Annetta, Minogue, Holmes, & Cheng, 2009). Thus, we expect time-on-task during a learning activity to be related to learning outcomes.

In the context of the present field study, the course's instructor suspected that students failing to start the project until late in the semester harmed their own learning; that is, by spending less time working on the project, the instructor believed students would not learn as much from completing it in comparison with students who did spend more time working on it. Prior research on time-on-task, as described above, supports this view.

Hypothesis 1: Time-on-task will be positively related to learning outcomes.

Using Leaderboards to Increase Time-on-task

Given our goal of increasing student Time-on-task (i.e., the target behavior in our model), we next sought a gamification approach that would be likely to increase this behavior, identifying leaderboards. Leaderboards represent a meaningful combination of game attributes (as defined by Landers, in press) with the potential to increase Time-on-task (see, for example, Brousell, 2013). They are also quite common (see Cheong, Cheong, & Filippou, 2013; Dominguez et al., 2013; Farzan et al., 2008; Hamari et al., 2014). Leaderboards represent a combination of three of the nine game attribute categories that Landers theorized to be valuable when gamifying learning: conflict/challenge, rules/goals, and assessment. We next explore the specific mechanisms by which each of these attributes should increase Time-on-task.

Conflict/Challenge. Landers (in press) defines the conflict/challenge game attribute as "problems provided to learners, including both the nature of difficulty of those problems" (p. XX). In education, classes are typically implicitly built upon the idea of competition as a means of motivating learners to act. Rewards (e.g., grades, degree completion) are typically based upon both a student's relative standing among peers and upon achievement of a set level of mastery desired by the instructor (Michaels, 1977). In the first case, students are competing with each other; in the second, students are competing with their instructor's vision of the perfectly taught student emerging from the classroom. In both cases, this sense of competition is intended to encourage the student to engage with the material on their own time and apply it to their lives to a greater degree than they would if not motivated in such a way.

When such competition is used and competitors are of roughly equal ability, success at the competition becomes contingent upon effort (Slavin, 1980). Under these circumstances, conflict has the maximum beneficial effect. If competitors are not of equal ability, less skilled competitors may become less motivated over time after repeated failures while more skilled competitors may become less motivated as the challenge diminishes. Thus, the relationship between conflict and effort is curvilinear, with an ideal level of conflict at a middle ground. The level of challenge presented by a learning activity has also been linked with time-on-task in a similar pattern; when students are frustrated or unchallenged, they spend less time-on-task than when the difficulty level is matched to their ability level (Treptow, Burns, & McComas, 2007).

Given this research, leaderboards should be designed to ensure that all learners have roughly equal chances of placing on those leaderboards, given equal amounts of effort. This will maximize the impact of leaderboards on time-on-task. This can be done by setting the leaderboard challenges adjacent to learning objectives. For example, in a course on English literature, one learning objective might be to understand the influence of Chaucer on subsequent writers in his time period. A leaderboard used in this course should not reward understanding of Chaucer, because such a leaderboard would over-reward students who already met this objective before the course began and under-reward students who were struggling with the idea despite putting in a great deal of effort. Instead, the leaderboard should be targeted at objectives that increase time-on-task. In this case, the instructor might ask students to log into a website each time they finished reading an optional work of Chaucer, and the number of works read so far might be displayed on a leaderboard. This approach gives all students an equal opportunity to appear on the leaderboard and broadly increases student motivation to engage with the Chaucer material while out of the classroom.

Rules/Goals. Landers defines the rules/goals game attribute as "clearly defined rules, goals, and information on progress toward those goals, provided to the player" (p. XX). Goal-setting theory is a valuable framework for understanding the potential effects of this attribute (Landers, Bauer, Callan, & Armstrong, 2015; Landers, Callan, & Bauer, Provisionally Accepted). In brief, this theory describes how goals most effectively motivate individuals to act when those goals are specific, measurable, achievable, realistic, and time-bound (SMART; Doran, 1981; Locke & Latham, 1990). A leaderboard can provide SMART goals if designed with learner abilities in mind, but it is important to note that leaderboard goals are not by definition SMART. For example, consider a leaderboard with the following item: *Completes a High-Quality Project*. Although it might be assumed that individuals at the top of the leaderboard would have completed higher quality projects than those further down the leaderboard, it is not clear what *high quality* means, making this an unspecific, unmeasurable goal. If the instructor has not provided sufficient tools or training to create a *high-quality* project, the goal may not be achievable. For specific learners, the goal may not be realistic if they do not possess sufficient self-efficacy to believe that they can create a *high-quality* project. Finally, although the project may have a natural conclusion point (e.g., the end of the course), the leaderboard goal is not similarly time-bound. It is unclear if the leaderboard is assessed continuously, at specific time points, or only at the conclusion of the project. All of these concerns must be addressed for this leaderboard goal to motivate learners to act, and a similar reasoning process is necessary to ensure every goal specified on the leaderboard is SMART.

Assessment. Landers (in press) defines the assessment game element as "the method by which accomplishment and game progress are tracked" (p. XX). In a leaderboard, assessment refers to the system by which accomplishment of rules/goals is formally recognized. For example, rankings of top players or point values might be displayed. Assessment thus forms the basis of social meaning created from the accomplishment of a leaderboard's rules/goals; without assessment, a leaderboard is only an unordered list of names and tasks.

Given this prior research suggesting the ability of the game elements utilized by leaderboards to increase their time-on-task, we next propose a hypothesis describing this relationship.

> **Hypothesis 2:** Learners randomly assigned to a project gamified with leaderboards will spend more time-on-task than learners randomly assigned to an otherwise identical project lacking leaderboards.

Obtaining support for these two hypotheses would demonstrate that gamification can cause changes in learner behavior and that learner behavior is correlated with learning outcomes. However, these hypotheses do not provide a test of the theory of gamified instruction. To do this, we must test both relationships simultaneously as a mediating effect. Essentially, we propose that the reason leaderboards improve learning outcomes is because leaderboards increase time-on-task, which in turn increases learning outcomes.

> **Hypothesis 3:** The relationship between leaderboard implementation and learning outcomes will be mediated by time-on-task.

Method

Participants

Study participants were members of an online upper-division industrial/organizational psychology course at a large U.S. east coast university. Students were primarily categorized as distance students, and thus were physically distant from each other without much opportunity to interact. One hundred nine students were randomly assigned to gamified and control conditions. Of those, 86 provided consent for their data to be used for research purposes ($n_{\text{Gamified}} = 42$, $n_{\text{Control}} = 44$), and 64 completed an end-of-course survey for extra credit in the course to provide additional demographic information. A chi-square test of independence indicated that the decision to participate in research did not differ by condition ($\chi^2 = .00$, $p > .05$), providing evidence against a participation bias. Reported ages of participants completing the demographics questionnaire ranged from 20 to 58, with a mean of 32.4. Reported grade point average of participants ranged from 1.8 to 4.0, with a mean of 3.19. Participants were 75% female. Majors were diverse: 27% Psychology majors, 22% in the Business School, 17% Engineering majors, 10% Science majors (other than Psychology), 9% Education majors, and 15% other majors or undeclared. Sixty-three percent of participants held a part-time job during the course, 22% were full-time, and 15% did not have a job. In general, the participants were fairly diverse in experience and qualifications, which reflected the general population of the university.

Materials

The course project used as the focus of this study required students to browse to a web-based wiki created with the open source MEDIAWIKI software platform. Students were assigned a topic as a research focus, and they were expected to write a wiki page exploring that topic over the course of a semester. These topics were related to course content, but not covered anywhere in course lecture videos or the textbook. Successful projects integrated information found via independent research with course content to meaningfully address the assigned topic. Completion of the project therefore required students to conduct independent research on a unique applied topic using university library resources and general web searches. On each wiki, two students were assigned to each topic. Because the subject of the course was industrial/organizational psychology (i.e., psychological science applied to the workplace) in a U.S. setting, these topics included *Team Failure Due to Groupthink: The Disaster of the Space Shuttle Challenger, Important Supreme Court Cases Related to the Civil Rights Act of 1964*, and *Practical Strategies for Being a Just Employer*, among others. A rubric was also provided emphasizing Formatting, Content, Visuals, and Citations as performance dimensions on the project (see the "Measures" section for definitions).

Measures

Time-on-task. To capture the amount of time spent by each student on their wiki project, the total number of edits made by each student on their project entry was recorded. This behavioral measure was captured objectively and without measurement error automatically by MEDIAWIKI and served as the best available proxy of time-on-task that did not rely upon self-report. Specifically, the number obtained from the MEDIAWIKI was the number of unique times each student saved the wiki page containing that student's project.

Academic performance. Based upon the assignment rubric for the project, we identified four dimensions of performance for the project as the outcome of this study: Formatting (professional and formatted appropriately), Content (high-quality ideas and topic coverage), Visuals (creativity and use of wiki design features), and Citations (appropriately citing ideas, text, and images used). Each category ranged on a scale from 1 to 10. Both authors of this article rated all 86 wiki entries in random orders, blind to experimental condition. The means of our two ratings were then used in analyses. To determine inter-rater reliability, we calculated intra-class correlations (ICC). ICC(2,2) was computed for each dimension (.86, .90, .93, and .98, respectively), which indicated high inter-rater reliability.

Procedure

To isolate the effects of gamification from confounds, and to enable conclusions about causality regarding the gamification intervention, the study utilized an experimental design. In the second week of the course, all students enrolled were randomly assigned to complete a semester-long course project on either a gamified or control wiki. Their assigned wiki was the only option visible to them from the course management system, and they were never told that the other half of the class had a different wiki. All inter-student communication tools other than the wiki itself were disabled so as to discourage any communication that might reveal the presence of two conditions, consistent with Cooke's (as cited in Wolfe & Crookall, 1998) recommendation that students remain blind to their experimental condition. In both wikis, students were paired with another student writing on the same topic and encouraged to compare their progress with that of the other student assigned to their topic.

Control and experimental wikis were identical except for the addition of two leaderboards to the experimental wiki. In the gamification condition, a leaderboard was added listing 10 action goals (see Figure 3) and awarded students for being the first, second, or third to accomplish each task. If a student completed a goal in one of these places, their name and a link to their wiki user profile would be added, making their accomplishment visible to anyone who opened the wiki. On the secondary leaderboard, each student's wiki username appeared aside their paired student's username, separated by a "vs.," and accompanied by two point totals. Each week, the course teaching assistant compared the two articles in each pair and awarded a point to whichever student would have earned a higher grade given the quality of their entry at that time, based upon the rubric provided. That student's username was then highlighted in bold to indicate that the student was winning. This was done for each of the 10 weeks that the project was open. This created both long-term high-stakes conflict (only three winners possible) and short-term low-stakes conflict (points awarded every week). It also counterbalanced focal topic difficulty across conditions (the same topics were covered within each condition, and each topic was covered by two students within each condition) so as to avoid confounds due to project difficulty.

Although all goals were project-related and achieving leaderboard goals was likely to result in better projects, appearing on the leaderboard did not itself directly

	1st Place	2nd Place	3rd Place
First Three to Create Their Entry			
First Three to Create a Profile			
First Three to Upload an Image to their Profile			
First Three to Upload an Image to their Article			
First Three to Reach 500 Words of Content			
First Three to Reach 1,000 Words of Content			
First Three to Reach 2,000 Words of Content			
First Three to Reach 5,000 Words of Content			
First Three to Edit Someone Else's Profile	edited	edited	edited
First Three to Edit Someone Else's Entry	edited	edited	edited
Most Edits as of March 1			
Most Edits as of April 1			
Most Edits as of the Due Date			
Best Profile as of March 1			
Best Profile as of April 1			
Best Profile as of the Due Date			
Best Entry as of March 1			
Best Entry as of April 1			
Best Entry as of the Due Date			

Figure 3. Wiki-based leaderboard illustrating provided goals.

influence grades. It was possible for a student to earn the maximum possible grade on the project without ever appearing on the leaderboard, and a student could also appear on the leaderboard, but still earn a low grade.

To minimize any differential treatment (and thus threats to internal validity) beyond the inclusion of the gamification elements themselves between conditions, and to maximize the generalizability of study results regardless of the specific implementation decisions made by any particular instructor, the leaderboards were also never discussed in any course materials outside of the wiki itself, including course announcements. The extent of the intervention was to place these two leaderboards in the wiki

Table 1. Descriptive Statistics and Correlation Matrix of Study Variables (n = 86).

Outcome	*M*	*SD*	Condition	Time-on-task	Format	Content	Visuals	References
Condition	0.49	0.50	—					
Time-on-task	55.86	42.77	.334*	—				
Formatting	5.72	2.43	−.032	.343*	(.862)			
Content	6.06	2.45	.063	.390*	.910*	(.896)		
Visuals	5.65	2.47	.167	.482*	.853*	.905*	(.933)	
References	6.92	4.72	.006	.474*	.704*	.764*	.729*	(.979)

Note. Condition was coded as 0 = control, 1 = gamification; ICC(2,2) inter-rater reliabilities appear on the diagonal of the correlation matrix. ICC = Intra-Class Correlations.
*p < .01.

Table 2. Ordinary Least Squares Regressions of Outcomes on Time-on-Task (n = 86).

Outcome	Intercept	*b*	r^2	*t* statistic	*p* value
Formatting	4.626	0.019	.118	3.347	.001
Content	4.803	0.022	.152	3.879	.000
Visuals	4.083	0.028	.232	5.043	.000
Citations	3.983	0.052	.225	4.933	.000

project, with no additional explanation provided to students; the leaderboard was simply present or it was not.

Results

A summary of all project variables appears in Table 1.

Hypothesis 1

To test whether time-on-task was positively related to academic performance, we regressed each of the four learning outcomes on time-on-task. The results of these analyses appear in Table 2. For all four outcomes, time-on-task strongly predicted outcomes and was statistically significant. Variance in outcomes explained by time-on-task ranged from 11.8% to 23.2%, moderate to large effects. Hypothesis 1 was supported.

Hypothesis 2

To assess whether gamification led to greater time-on-task, we regressed time-on-task on gamification condition (dummy coded as 0 or 1). The slope was statistically significant, $b = 29.61$, $r = .33$, $r^2 = .12$, $t(84) = 3.40$, $p = .001$. Approximately 12% of the

Table 3. Tests of Time-on-Task as a Mediator of Gamification's Effect on Outcomes (n = 86).

Outcome	Point estimate	*SE*	95% CI LL	95% CI UL	*z* statistic	*p* value
Sobel tests						
Formatting	0.676	0.274	0.137	1.214	2.460	.014
Content	0.709	0.282	0.156	1.261	2.515	.012
Visuals	0.826	0.304	0.230	1.422	2.715	.007
Citations	1.752	0.620	0.537	2.967	2.827	.005
Bootstrapping						
Formatting	0.680	0.278	0.218	1.293		
Content	0.706	0.287	0.218	1.321		
Visuals	0.839	0.327	0.295	1.567		
Citations	1.664	0.533	0.674	2.798		

Note. CI = Confidence Interval; LL = Lower Limit; UL = Upper Limit.

variance in time-on-task was explained by the gamification condition, a moderate effect (Cohen, 1988). On average, students in the gamification condition made 29.61 more edits than those in the control condition. Because gamification was controlled experimentally, this provides strong evidence that the use of the leaderboard caused an increase in time-on-task. Hypothesis 2 was supported.

Hypothesis 3

To test whether the relationship between gamification condition and outcomes was mediated by time-on-task, one Sobel test and one bootstrapped confidence interval of the indirect effect were calculated for each outcome. Although both the Sobel test and bootstrapping provide statistical evidence of mediation, under certain conditions they may disagree, and if so, the cause of the disagreement should be identified (Preacher & Hayes, 2004). The results of these analyses appear in Table 3. For all outcomes, both Sobel tests and bootstrapping agreed and were statistically significant (i.e., confidence intervals did not include zero), indicating time-on-task as a consistent mediator of the relationship between gamification condition and outcome. Hypothesis 3 was supported.

Discussion

This article makes four primary contributions to the growing literature on gamification and serious games. First, we developed a specific theoretical rationale to guide the use of leaderboards to increase learner time-on-task. Two game elements common to leaderboards, rules/goals and conflict/challenge, were theorized as the primary drivers of these relationships. The rules/goals attribute category should accomplish this as predicted by goal-setting theory; leaderboard tasks must be specific, measurable, achievable, realistic, and time-bound. The conflict/challenge attribute category motivates students to defeat the challenges presented to them, but must be designed so that motivation and effort are the primary drivers of success (instead of ability). Without this design element, conflict/challenge could be demotivating. The assessment element is also used to provide a record of student accomplishment, which supports the effects of the other two elements.

Second, this study provides strong empirical support for the causal effect of gamifying a course project with leaderboards to improve time-on-task. By utilizing an experimental design with true random assignment of students to condition, one of a small set of valid approaches to demonstrating causality, we conclude that leaderboards can be successfully applied in education to influence learner behavior. However, the ultimate value of this behavioral change is still contingent on its role as a moderator or mediator in the specific gamification process modeled.

Third, this study provides strong empirical support for the indirect effect of gamification on learning through behavioral change, supporting the theory of gamified learning, with specific evidence supporting the time-on-task construct as a valid behavioral target for gamification. From our statistical tests, we first conclude that the behavior targeted by gamification in this study, time-on-task, predicts learning outcomes. Given our theoretical rationale supporting the use of game elements and the empirical support found for the leaderboard intervention, we conclude that the conflict/challenge, rules/goals, and assessment game elements used in combination as a leaderboard causally, but indirectly affect learning outcomes through the time-on-task mediator in this study.

Fourth, the research design of this study meets the requirements presented by Cooke (as cited in Wolfe & Crookall, 1998) as ideal for rigorous educational research. All groups were identically composed, all material was delivered within a single semester, all groups were created randomly, groups were unaware of the experiment, the innovation's content was nested within the course's subject matter, and a wide range of scores were observed. This was despite Wolfe and Crookall's assertion that such a design was impossible to achieve in serious games research. We thus present this study as an example of a rigorous quantitative research design as recommended by Cooke examining the application of game principles. We recommend continued quantitative research developing a basic-science level understanding of game elements in both gamification and serious games. Researchers should incorporate careful consideration of design to carefully isolate the effect or effects of interest from other confounding factors common in instructional settings.

Limitations

We identified four primary limitations to this study. First, although we identified a causal path from gamification elements to outcomes via time-on-task, some behaviors and attitudes are likely more effective than others. Those suggested by Deterding, Sicart, Nacke, O'Hara, and Dixon (2011), namely, user experience and user engagement, are especially worth exploration. Many other behaviors and attitudes that impact learning and could be caused by gamification likely exist beyond those listed here, and this research should not be interpreted as a call to focus solely upon time-on-task. Future researchers should work to develop taxonomy of learning-related behaviors that game elements can be used to successfully change.

Second, we examined gamification of learning within the context of a U.S. undergraduate student sample. Although this is a natural setting for a study of gamification in education, the motivational forces of employees conducting mandatory training may be quite different. For example, this research may better generalize to mandatory workplace training than to voluntary workplace training. Additional research on game elements should be conducted with employee samples and any resulting differences explored. If generalizing to education, our sample was a relatively diverse collection of students; a more homogeneous collection of students might have strengthened the relationship between time-on-task and performance. In addition, the use of a U.S. sample may have increased the willingness of students to engage in competition in the classroom in comparison with students in other cultures. Future research should investigate potential cultural differences in the effectiveness of this and other gamification interventions.

Third, the experimental condition did not correlate strongly with learning outcomes (see Table 1); that is, if we had not examined the mediation effect explicitly, gamification would not have appeared to have affected it. A direct relationship between antecedent and outcome is not required for mediation (Preacher & Hayes, 2004); however, its absence indicates that the effect of gamification is subtle and indirect, entirely through the mediating construct time-on-task. We thus recommend future gamification researchers to carefully identify target behaviors/attitudes and ensure their appropriate measurement. Because the effect of gamification on learning is indirect, failure to measure the mediator may lead to erroneous conclusions.

Fourth, the theory of gamified learning on which this article is based (Landers, in press) makes a case that debriefing is not necessary in gamification, and we omitted it here accordingly. However, debriefing of serious games might itself be gamified. For example, in the framework of the theory of gamified learning, debriefing could be itself considered a type of instructional content; a gamification intervention might be used to increase motivation to participate actively in debriefing. Debriefing might also be used as a way to verify the psychological processes involved in gamification's impact; although we have implemented game elements as an intervention, additional psychological mediators may further explain observed effects. Further research should examine both the potential role of debriefing in gamification and the role of gamification in debriefing.

Practical Recommendations and Conclusions

Given our four major contributions outlined above, we make several practical recommendations to those seeking to implement gamification. First, practitioners seeking to adopt gamification should identify the specific behavioral target of that gamification. In our study, the target was to improve time-on-task. If we had only examined gamification and learning outcomes, we would have been misled in our conclusions. Second, practitioners should identify the mechanism by which their targeted behavior should affect learning. Two processes are possible: (a) The behavior strengthens the relationship between instructional effectiveness and outcomes and/or (b) the behavior causally affects learning directly. In our study, we chose to target time-on-task because of its previously identified direct relationship with learning outcomes, suggesting a mediating process. Third, practitioners should ensure no contextual or situational factors affect the causal paths identified. If the behavior is intended to causally affect learning, practitioners should ensure that both the game element affects the behavior and that the behavior affects learning. If the behavior is intended to increase the effectiveness of pre-existing instruction, practitioners should ensure that both the game element affects the behavior and that the behavior and instruction work synergistically (i.e., a two-way interaction is present). Fourth, we recommend rigorous evaluative techniques appropriate to the type of training/instruction conducted (see Kraiger, Ford, & Salas, 1993; Kuncel & Campbell, 2001; Landers & Callan, 2012).

We also provide a note of caution to instructional designers considering the use of leaderboards and gamification in general. The nature of the learning task to be gamified and the specific game elements to be used must be carefully considered. This article should not be interpreted as a general recommendation to add leaderboards whenever student motivation is low. Instead, processes that could improve learning (such as increased time-on-task) must be identified, and those processes must be targeted by gamification interventions in order to affect learning indirectly. Gamification interventions targeted directly at learning are less likely to be successful and may be harmful for many learners (Callan, Bauer, & Landers, 2015). Even when targeted to affect learning indirectly, leaderboards may be the best approach in some cases, whereas in others, collaborative approaches (i.e., the human interaction game element), the incorporation of a narrative (i.e., the game fiction element), or the use of other game elements may be more appropriate and effective. Even when leaderboards are the best approach, many minor design decisions are involved in the implementation of leaderboards that may influence the impact of those leaderboards that have not yet been explored.

In summary, we have provided a rigorous empirical test of the mediational process of the theory of gamified learning. Gamification can be used successfully to affect targeted learning-related behaviors, and these behaviors are linked with improved learning outcomes. To our knowledge, this study is the first empirical work demonstrating this effect. Thus, we hope that this study will provide a strong basis for continued examination of the theory of gamified learning (Landers, in press) by directing future research to integrate Bedwell and colleagues' (2012) taxonomy of game elements into their research, in order to test the value of those elements.

Author Contributions

Both authors contributed to this article, in content and in form. Both authors designed the experiment and coded the academic achievement variable. RNL ran the course used as the experimental setting and executed the study protocol. RNL also conducted the statistical analyses and wrote the first draft of the manuscript. Both authors edited the manuscript.

Acknowledgments

The authors thank seven anonymous reviewers for their comments on earlier versions of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, *53*, 74-85.

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: An Interdisciplinary Journal*, *43*, 729-760.

Brousell, L. (2013). *How gamification reshapes corporate training*. Retrieved from http://www.cio.com/article/728268/How_Gamification_Reshapes_Corporate_Training

Brown, K. G. (2001). Using computers to deliver training: Which employees learn and why? *Personnel Psychology*, *54*, 271-296.

Callan, R. C., Bauer, K. N., & Landers, R. N. (2015). How to avoid the dark side of gamification: Ten business scenarios and their unintended consequences. In T. Reiners & L. Wood (Eds.), *Gamification in education and business* (pp. 553-568). New York, NY: Springer.

Cheong, C., Cheong, F., & Filippou, J. (2013, June). *Quick quiz: A gamified approach for enhancing learning*. In Proceedings of Pacific Asia Conference on Information Systems, Jeju Island, South Korea.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011, May). *Gamification: Toward a definition*. Proceedings of the CHI 2011 Gamification Workshop, Vancouver, British Columbia, Canada.

Dominguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernandez-Sanz, L., Pages, C., & Martinez-Herraiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, *63*, 380-392.

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*, *70*(11), 35-36.

Ericsson, K. W., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.

Farzan, R., DiMicco, J. M., Millen, D. R., Brownholtz, B., Geyer, W., & Dugan, C. (2008, April). *When the experiment is over: Deploying an incentive system to all the users*. Symposium on Persuasive Technology, Aberdeen, Scotland.

Goldstein, I. L. (1993). *Training in organizations: Needs assessment, development, and evaluation*. Monterey, CA: Brooks/Cole.

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January 6-9). *Does gamification work? A literature review of empirical studies on gamification*. In Proceedings of the 47th Hawaii International Conference on System Sciences, Waikoloa, HI.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, *74*, 657-690.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, *78*, 311-328.

Kuncel, N. R., & Campbell, J. P. (2001). Individual and team training. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work & organizational psychology* (Vol. 2, pp. 278-312). London, England: SAGE.

Landers, R. N. (in press). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming: An Interdisciplinary Journal*.

Landers, R. N., Bauer, K. N., Callan, R. C., & Armstrong, M. B. (2015). Psychological theory and the gamification of learning. In T. Reiners & L. Wood (Eds.), *Gamification in education and business* (pp. 165-186). New York, NY: Springer.

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. 399-424). Surrey, UK: Springer.

Landers, R. N., & Callan, R. C. (2012). Training evaluation in virtual worlds: Development of a model. *Journal of Virtual Worlds Research*, *5*(3). Retrieved from https://journals.tdl.org/jvwr/index.php/jvwr/article/view/6335

Landers, R. N., Callan, R. C., & Bauer, K. N. (Provisionally Accepted). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior*.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice Hall.

Michaels, J. W. (1977). Classroom reward structures and academic performance. *Review of Educational Research*, *47*, 87-98.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717-731.

Pychyl, T. A., Lee, J. M., Thibodeau, R., & Blunt, A. (2000). Five days of emotion: An experience sampling study of undergraduate student procrastination. *Journal of Social Behavior and Personality*, *15*, 239-254.

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, *50*, 315-342.

Treptow, M. A., Burns, M. K., & McComas, J. J. (2007). Reading at the frustration, instructional, and independent levels: The effects on students' reading comprehension and time on task. *School Psychology Review*, *36*, 159-166.

Wolfe, J., & Crookall, D. (1998). Developing a scientific knowledge of simulation/gaming. *Simulation & Gaming: An Interdisciplinary Journal*, *29*, 7-19.

Author Biographies

Richard N. Landers, PhD, is an assistant professor of industrial/organizational psychology at Old Dominion University. His research program focuses upon improving the use of Internet

technologies in talent management, especially the measurement of knowledge, skills, and abilities; the selection of employees using innovative technologies; and learning conducted via the Internet. He has been a video game enthusiast since 1984.

Contact: rnlanders@odu.edu.

Amy K. Landers, PhD, is a lecturer and the director of the online psychology major at Old Dominion University. Her research interests include determining the effectiveness of teaching techniques for online education, improving outcomes for students enrolled in online courses, and identifying ways to improve teaching in an online environment.

Contact: alanders@odu.edu.