



Gamification of task performance with leaderboards: A goal setting experiment



Richard N. Landers ^{a,*}, Kristina N. Bauer ^b, Rachel C. Callan ^a

^a Old Dominion University, USA

^b Illinois Institute of Technology, USA

ARTICLE INFO

Article history:

Received 11 June 2014

Received in revised form

10 June 2015

Accepted 11 August 2015

Available online 28 August 2015

Keywords:

Gamification

Goals

Goal-setting theory

Leaderboard

Quantitative

Experiment

ABSTRACT

The use of leaderboards is a common approach to the gamification of employee performance, but little is known about the specific mechanisms and mediating processes by which leaderboards actually affect employee behavior. Given the lack of research in this domain, this study proposes goal-setting theory, one of the most well-established motivational theories in psychology, as a framework by which to understand these effects. In this study, a classic brainstorming task is gamified with a leaderboard in order to explore this. Participants were randomly assigned to four classic levels of goal-setting (do-your-best, easy, difficult and impossible goals) plus a leaderboard populated with initials and scores representing identical goal-setting conditions. The presence of a leaderboard was successful in motivating participants to performance levels similar to that of difficult and impossible goal-setting, suggesting participants implicitly set goals at or near the top of the leaderboard without any prompting to do so. Goal commitment, a common individual difference moderator in goal-setting theory, was also assessed and behaved similarly in the presence of the leaderboard as when traditional goals were provided. From these results, we conclude that goal-setting theory is valuable to understand the success of leaderboards, and we recommend further exploration of existing psychological theories, including goal-setting, to better explain the effects of gamification.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

In the organizational context, gamification is a promising avenue by which to increase employee task performance (i.e., in-role behavior), one dimension of individual work performance (Koopmans et al., 2012). By directing and rewarding employee attention to particular focal tasks through goal setting, performance can be improved (Locke & Latham, 2002). Although there are at least forty years of research describing how motivational interventions can be used to alter task performance with goals (e.g., Latham & Baldes, 1975), little empirical research is available exploring the use of gamification as such an intervention (Hamari, Koivisto, & Pakkanen, 2014), despite the use of goals as a key component of most gamification interventions.

The specific way that goals are implemented differs between traditional goal-setting efforts and gamification. In traditional goal-setting, a single specific goal (or group of goals) is set for an

employee to achieve. There is even something approaching scientific consensus as to which types of goals are the most motivating: those which are specific, measurable, attainable, realistic, and time-bound (SMART; Moskowitz & Grant, 2009). When a goal is provided to employees, employees can choose whether or not to pursue that goal. Within gamification efforts, the use of points and leaderboards, two of the most common types of gamification, can both be interpreted as applications of non-optimal goal-setting. When points alone are implemented to track achievement, there is no specific goal to pursue. Instead, the employee decides what quantity of points is worth pursuing. Similarly, a leaderboard presents many possible goals representing the prior performance levels of all those appearing on it. Thus, both points and leaderboards require employees to set their own goals, which is a more subtle attempt at behavior modification than traditional goal-setting interventions. In most organizations, supervisors are unlikely to make placement on a leaderboard a job requirement; however, the mere presence of the leaderboard may imply this.

Given this, we contend that because the effectiveness of goal setting has been extensively explored in psychology, it will serve as an effective basis for understanding the effectiveness of

* Corresponding author. 250 Mills Godwin Building, Norfolk, VA 23529, USA.
E-mail address: rnlanders@odu.edu (R.N. Landers).

leaderboards and the point totals they present. Experimental research in particular is critical on this point, because it is the only approach that can provide causal evidence of impact by managerial efforts on employee and ultimately organizational performance (Pedhazur & Schmelkin, 1991). Thus, the purpose of this paper is to respond to calls by Hamari (2013) and Landers and Callan (2011) by experimentally examining the effectiveness of goal-setting theory to explain observed changes in task performance resulting from a leaderboard intervention.

1.1. Effects of gamifying with leaderboards

Preliminary empirical evidence is available for the effectiveness of leaderboards to alter a variety of outcomes. In the learning context, Dominguez et al. (2013) examined differences between students quasi-experimentally assigned to view leaderboards ranking badge attainment, finding that students viewing the leaderboard scored higher on some assignments but lower on others. Eickhoff, Harris, de Vries and Srinivasan (2012) included leaderboards as part of a larger game to encourage flow states. Although the effects of leaderboards were not isolated in their research design, the researchers concluded that the inclusion of leaderboards “was a moderate success” (p. 9) in encouraging competition. In the context of web-based social communities, Farzan, DiMicco, Millen, et al. (2008) added leaderboards as part of a larger incentive system to a website, finding increases in user contributions four weeks after deployment. In a crowdsourcing application for conversational modeling, Halan, Rossen, Cendan, and Lok (2010) incorporated leaderboards alongside narratives and deadlines to increase participation, finding that participation increased but that those participating were more likely to behave unrealistically while conversing.

It is difficult to conclude in any of these studies that leaderboards, specifically, caused the observed effects reported, because leaderboards are rarely experimentally isolated as a gamification technique. When leaderboards are included in experimental conditions with other game elements, like badges or narrative, the presence of those additional game elements or any interaction between leaderboards and those game elements may actually be driving experimentally observed differences. Such designs may speak to the ability of gamification to affect outcomes generally, but they do not help to explain the effects of leaderboards specifically. From the standpoint of practice, the current empirical gamification literature can therefore provide no recommendations on the causal impact of leaderboards when implemented in a novel situation. Thus, a major goal of the present study is to isolate leaderboards to examine their causal effects in order to provide such recommendations.

1.2. Goal-setting theory

Goal-setting theory was originally developed by Locke (1968), who proposed that people will be motivated to strive towards goals. This approach is effective due to the psychological process of self-regulation (Latham & Locke, 1991), which acts as a mediator (i.e., intermediary causal process) between set goals and performance (Kanfer & Ackerman, 1989). Self-regulation can be defined as the modification of thought, affect, and behavior (Karoly, 1993). The goal provides the individual a measure for “excellent” performance against which to judge his or her own performance. The individual can subsequently alter his or her behavior in order to reduce the discrepancy between the performance and the goal (Latham & Locke, 1991). Goal-setting interventions are considered to be among the most powerful motivational interventions, found to be effective across many situations and tasks (Locke & Latham,

2002). When used as an intervention, leaderboards likely perform similarly to classic goal-setting interventions, because leaderboards provide the user with several potential goals. The user should be motivated to reach one of these goals and regulate his or her behavior by reducing the discrepancy between the desired goal from the leaderboard and actual performance until the given goal is met.

The effects of goal setting on self-regulation are consistent across a range of outcomes (Locke & Latham, 2002). For example, Frayne and Geringer (2000) demonstrated that job performance can be improved through self-management training, which included self-monitoring, goal-setting, and relapse prevention components. Over 12 weeks, employees were able to sustain job performance improvements when compared to the control group who received no such training. This relationship has also been observed in salespeople, where self-regulatory tactics, including goal-setting, were found to fully mediate the relationship between goal orientation and sales performance (VandeWalle, Brown, Cron, & Slocum Jr, 1999). This suggests that those who are seen as having advantageous goal orientations effectively use self-regulation to set goals and manage their performance. Similar results were found when investigating self-regulation and goal-setting among students and the effects of class performance (Lee, Sheldon, & Turban, 2003). The relationship between self-regulation and goal-setting was also supported in a comprehensive meta-analysis that explored these effects across 85 studies (113 unique samples), spanning numerous contexts, nationalities, and age groups (Burnette, O’Boyle, VanEpps, Pollack, & Finkel, 2013). Together, this evidence suggests that the relationship between goal-setting and self-regulation is robust and these relationships should be expected in traditional and gamification contexts alike.

Latham and Baldes’ (1975) classic logger study provides an illustrative example of how goal-setting interventions can be used to improve performance, testing Locke’s (1968) original goal-setting theory. In their study, Latham and Baldes demonstrated that goal-setting was an effective method for improving the task performance of logging teams who had regularly been loading their trucks far short of their maximum capacity (approximately 60%). This was brought to the attention of management and union leaders, who agreed that a goal of 94% of the maximum capacity was a difficult but attainable goal for which the teams should strive. The loggers were paid hourly and told that the goal was part of the experiment; they would not receive any reward for meeting the goal nor would they be punished if they failed to reach it. After this goal-setting intervention, the truck weights were monitored for nine months. In the month the 94% goal was introduced, a spike in performance occurred in which the truck weights averaged just above 80% of capacity. Performance continued to climb to just above 90% of capacity, where it plateaued for the last 6 months of the study. Latham and Baldes (1975) argued that before the intervention, the loggers were essentially operating under a “do your best” goal, which when compared to the difficult, specific goal of 94% was less effective at motivating performance. In this study, the loggers monitored their own efforts in order to meet the goal, which caused the loggers to effectively utilize self-regulation.

As first seen in Latham and Baldes (1975), different goals have different effects on performance, and this has been demonstrated repeatedly in the related literature (Locke & Latham, 2002). Meta-analysis has demonstrated that there is a linear relationship between goal difficulty and performance, with estimated effect sizes ranging from $d = .52$ and $d = .82$ (Locke & Latham, 1990). From this evidence, we expect that when presented easier goals, an individual’s performance will be lower than his or her performance when given more difficult goals. Given the support for specific,

difficult goals in the literature, [Locke and Latham \(1990\)](#) also compared this to a common organizational practice of asking individuals to “do their best”. The meta-analytic effect sizes ranged from $d = .42$ to $d = .80$, which suggests that we can expect individual performance to suffer under a do-your-best goal when compared to performance under a specific, difficult goal. Impossible goals have also been investigated, with [Locke \(1982\)](#) demonstrating that even under impossible goals, performance does not suffer when compared to difficult goals and may even improve slightly, with the only noticeable difference being that some participants reported shifting their focus from reaching the goal to simply getting close to the goal. Given these findings, a pattern emerges, such that we can expect performance to differ across different goal-setting conditions with specific, difficult goals and impossible goals resulting in superior performance when compared to easy and do-your-best conditions.

In contrast to traditional goal setting interventions, leaderboards typically provide multiple optional goals and do not provide specific instructions on which of the presented goals is a worthwhile target. Leaderboards may be simply added to a task, and the performance of individuals may be tallied on that leaderboard. Across contexts, the use of leaderboards is one of the most common gamification interventions, and users vary in their motivations to achieve a place on such leaderboards ([Hamari, Koivisto, & Pakkanen, 2014](#); [Hamari, Koivisto, & Sarsa, 2014](#)). For example, [Landers and Landers \(2015\)](#) experimentally examined the addition of a leaderboard to a university course assignment, finding that its presence was positively related to overall time spent working on the assignment. Although some people exposed to leaderboards appear to strive to reach the top rankings, others appear satisfied to merely place ([Hamari, Koivisto, & Sarsa, 2014](#)). The degree to which goal-setting alone is responsible for such behaviors is currently unknown ([Landers, Bauer, Callan, & Armstrong, 2015](#)). Yet, we posit that individuals will strive to place near the top of the leaderboard, a position that is often difficult to attain. Thus, we expect that the leaderboard is effective because it serves as a difficult goal. Consistent with goal-setting theory, we hypothesize the following.

Hypothesis 1. *The leaderboard will function similarly to a difficult goal. Specifically, participants in the leaderboard condition should outperform participants in easy or do-your-best goal conditions.*

1.3. The role of goal commitment in goal-setting theory

Although goal setting interventions are generally effective, goal commitment has been found to be an important moderator in the relationship between goals and performance ([Latham & Locke, 1991](#)). According to [Latham and Locke \(1991\)](#), commitment acts as a moderator such that for those with high commitment, the linear relationship between goal difficulty and performance described above is observed; however, for those with low commitment, there is no relationship between goal level and performance. [Erez and Zidon's \(1984\)](#) results suggest that those with lower levels of commitment are more likely to reject difficult goals and replace them with easier goals, which helps explain the mechanism behind this moderating relationship. For those with stronger goal commitment, performance remained high even under impossible goal conditions, presumably because they continued to strive towards the impossible goals rather than revise their goals downward. In short, employees must commit themselves to a goal for that goal to be an effective motivator, and exceptionally committed employees will even strive toward that goal when it is unrealistic. We expect that the use of leaderboards will result in a similar process such that some individuals will

commit strongly to achieving the top spot on the leaderboard, whereas others will commit to less difficult goals, such as beating the lowest score on the leaderboard or beating a middle score.

Meta-analysis has also demonstrated the importance of goal commitment across a variety of contexts and outcomes. [Klein, Wesson, Hollenbeck, and Alge \(1999\)](#) performed a comprehensive meta-analysis to investigate the effects of goal commitment and found that across 66 independent samples, commitment was able to explain 5.3% ($r_c = .23$) of the variance in performance. These studies varied by task type, performance metrics, and goal origin, so these results can be considered to be generalizable across a variety of contexts and outcomes. As such, we contend that in the context of a gamified task, commitment will continue to play an important role in performance, exerting both direct and moderating effects, as described by [Latham and Locke \(1991\)](#). Goal commitment likely explains incremental variance in the success of goal setting across a variety of contexts and outcomes. This was demonstrated meta-analytically by [Klein et al. \(1999\)](#), who found that goal level moderated the relationship between goal commitment and performance. Specifically, goal commitment accounted for 12.3% ($r_c = .23$) of variance in task performance for goals of high difficulty but only 4.0% ($r_c = .20$) and 3.2% ($r_c = .18$) for goals of medium and low difficulty, respectively.

Goal commitment may function somewhat differently in the presence of leaderboards, because the leaderboard provides an opportunity for individuals to choose their own goals. According to [Latham and Locke \(1991\)](#), goal choice influences goal commitment because a person chooses a goal based on his or her expectation that he or she can reach that goal. In the case of a leaderboard, a person could choose to meet or surpass any particular score on the leaderboard as his or her goal. Such goals may even be revised upward as the person scores higher. The absence of such commitment is potentially damaging; for example, [Hamari \(2013\)](#) suggested that a lack of goal commitment was a likely cause for the weak effect of the gamification intervention he tested, concluding that it should be measured in future research. Given this, we expect that goal commitment will function as a moderator of the leaderboard-performance relationship just as it is a moderator of traditional goal setting and present here the first empirical test of this variable in gamification, to our knowledge.

Hypothesis 2. *Goal commitment moderates the relationship between the use of leaderboards and task performance. Specifically, greater goal commitment will strengthen the effect of more difficult goals and the leaderboard.*

2. Method

2.1. Participants

A total of 339 participants from a university in the southeastern United States were recruited through the research subject pool. They received either course credit or extra credit for participation. The mean age of participants was 22.40 ($SD = 6.85$), and all participants were over 18 years of age. The sample was .3% ($N = 1$) high school students, 31.9% ($N = 108$) freshmen, 21.2% ($N = 72$) sophomores, 26.0% ($N = 88$) juniors, 17.4% ($N = 59$) seniors, 2.4% ($N = 8$) returning or non-traditional students, and three participants chose not to respond. With respect to gender, 76.7% ($N = 260$) were female, 22.7% ($N = 77$) were male, .3% ($N = 1$) was transgender, and one participant opted not to respond. There were 55.5% ($N = 188$) White, 22.7% ($N = 77$) Black, 7.4% ($N = 25$) Asian, .9% ($N = 3$) Native American, 1.8% ($N = 6$) Hawaiian or other Pacific Islander, 7.4% ($N = 25$) mixed race, 3.8% ($N = 13$) other ethnicity, and two participants opted not to respond.

2.2. Materials

2.2.1. Experimental task

The experimental task for this study was a 12-minute brainstorming session where participants were asked to list uses for a knife, which was adapted from the brainstorming task used by White, Kjelgaard, and Harkins (1995) to demonstrate goal-setting effects. Similar tasks have also been used to demonstrate the effects of goal type (Litchfield, Fan, & Brown, 2011), performance feedback and goals (Jung, Schneider, & Valacich, 2005), and competition on goal-setting effectiveness (Hinsz, 2005), among others. Brainstorming tasks in general have been identified as an ideal context for much goal-setting research related to task performance (Litchfield, 2008). Prior to completing the experimental task, participants were instructed that they would be asked to generate as many uses for an item as possible and to focus on quantity as opposed to quality. They were also told to separate each use with a comma (,) and a space so that the program could automatically count and display the number of uses generated. Participants saw the number of uses they generated in real time, which provided immediate feedback. If the participant wrote a use identical to one previously written, this number did not increase.

2.2.2. Goal conditions

There were four goal conditions: (1) do-your-best, (2) easy goal, (3) difficult goal, and (4) impossible goal. In the do-your-best condition, participants were told to list as many uses as possible. In the goal conditions, participants were asked to generate a specific number of uses: 15, 39, and 53 for the easy, difficult, and impossible conditions, respectively. These goal levels are based on prior research using the brainstorming task (e.g., Szymanski & Harkins, 1987; White et al., 1995) and represent approximately 1 SD below, 1 SD above, and 3 SDs above the mean for the easy, difficult, and impossible conditions, respectively.

2.2.3. Leaderboard

Instead of the simple count, participants in the leaderboard condition were shown a leaderboard during the experimental task with pre-established scores on it associated with randomly chosen initials to imply that these scores had been achieved by specific people at some point in the past. The first, third, and fifth place scores corresponded to the impossible, difficult, and easy goals given to those in the goal-setting conditions. However, the number of uses generated was scored as points such that participants earned 7 points for each use generated (i.e. 15, 39 and 53 uses became 105, 273 and 371 points, respectively). Two intermediate goals (halfway between easy and difficult, and halfway between difficult and impossible) were also added to increase the number of choices and to make the leaderboard appear more realistic. Participants automatically moved up in rank on the leaderboard as they typed more unique uses. As occurred in the goal conditions, repeated ideas were not awarded points.

2.2.4. Goal commitment

Goal commitment was measured with a 5-item scale adapted from Hollenbeck, Klein, O'Leary, and Wright (1989). Each item was rated on a 5-point scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). In the goal conditions, items referenced the assigned goal, whereas in the leaderboard condition, items referenced the leaderboard (e.g., the implicit goal). A sample item is, "I was strongly committed to [pursuing the goal/scoring highly on the leaderboard]." Reliability was above generally accepted cutoffs ($\alpha = .80$; Nunnally, 1978). A confirmatory factor analysis was run, and loadings ranged from .62 to .86 suggesting that items loaded in

the very good to excellent range (Comrey & Lee, 1992). The complete list of items can be found in the Appendix.

2.2.5. Manipulation check

To ensure participants were paying attention to the instructions they were provided, two items were used, which differed by condition. For participants in the goal conditions, participants were asked to identify the goal they were given from a list of four options, including an option that there was no goal. For participants in the leaderboard condition, participants were asked to identify the highest score on the leaderboard. Participants were provided a list of possible options, including the option that there was no leaderboard. Thus, the manipulation checks were designed to detect those participants paying attention to study instructions.

2.3. Procedure

Participants signed up for the completely online study through the research pool and were directed to a website with an informed consent. After electronically signing the informed consent, participants completed demographic questions. Then, they received instructions for the practice task and were given three minutes to generate as many uses for a box as possible. This practice task allowed participants to get used to brainstorming as well as to the online system. After participants completed the practice task, they were randomly assigned to one of five conditions (single-factor between-subjects experimental design): (1) do-your-best, (2) easy goal, (3) difficult goal, (4) impossible goal, and (5) leaderboard. All participants engaged in the 12-min brainstorming session and then completed the goal commitment measure and manipulation check.

3. Results

Data were first cleaned and inspected for outliers. Eighty participants failed the manipulation check and another 13 were identified as performance issues. Performance issues included taking longer than the allotted time, listing only synonymous uses on the performance task, and listing nonsense answers. Consistent with recent recommendations for identifying careless responding (Meade & Craig, 2012), we used Mahalanobis distance to identify careless responding on the goal commitment items. This analysis resulted in the removal of five participants; one additional participant who did not complete the goal commitment items was also removed. Thus, analyses were run on 240 participants. We also checked the normality of goal commitment and performance; both were normally distributed. The descriptive statistics for these variables broken down by condition are reported in Table 1.

Demographic information of this group was very similar. The mean age of participants was 22.76 ($SD = 7.32$), and all participants were over 18 years of age. There was 1 high school student, 76 freshmen, 51 sophomores, 62 juniors, 42 seniors, 7 returning or non-traditional students, and 1 participant chose not to respond. With respect to gender, there were 179 females, 60 males, and 1 transgender participant. There were 139 White, 49 Black, 14 Asian, 2 Native American, 6 Hawaiian or other Pacific Islander, 20 mixed race, and 9 other ethnicity participants; 1 participant opted not to respond to this question. Overall, intercorrelations are presented in Table 2.

To examine the first hypothesis, a one-way ANOVA was run to determine whether the leaderboard condition outperformed the other four conditions. The Levene's test of equality of variance was significant, $F(4, 235) = 4.70, p = .001$. However, ANOVA is robust to moderate violations of homogeneity of variance and is a conservative test when groups with fewer people have smaller variance (Maxwell & Delaney, 2004). Therefore, we proceeded with a

Table 1
Descriptive statistics for performance and goal commitment by condition.

| Condition | Performance | | | | Goal commitment | | | | N |
|-----------------|-------------|-------|--------|---------------|-----------------|------|--------|------------|-----|
| | Min | Max | Median | M (SD) | Min | Max | Median | M (SD) | |
| Do Your Best | 8.00 | 65.00 | 20.00 | 23.84 (12.88) | 2.00 | 5.00 | 4.00 | 3.71 (.63) | 49 |
| Easy Goal | 5.00 | 48.00 | 19.00 | 22.10 (9.22) | 2.00 | 5.00 | 3.80 | 3.70 (.73) | 41 |
| Difficult Goal | 6.00 | 72.00 | 39.50 | 36.56 (14.35) | 1.40 | 5.00 | 3.80 | 3.75 (.71) | 66 |
| Impossible Goal | 8.00 | 75.00 | 42.00 | 41.55 (16.34) | 2.20 | 5.00 | 3.80 | 3.60 (.73) | 51 |
| Leaderboard | 9.00 | 88.00 | 33.00 | 37.52 (19.45) | 2.20 | 4.40 | 3.20 | 3.33 (.59) | 33 |
| Total | 5.00 | 88.00 | 30.00 | 32.68 (16.42) | 1.40 | 5.00 | 3.80 | 3.64 (.69) | 240 |

Table 2
Intercorrelations among study variables.

| | 1 | 2 | 3 | 4 | 5 |
|-------------------------------|---------|---------|---------|--------|--------|
| 1. Leaderboard vs. DYB | | | | | |
| 2. Leaderboard vs. Easy | -.23*** | | | | |
| 3. Leaderboard vs. Difficult | -.31*** | -.28*** | | | |
| 4. Leaderboard vs. Impossible | -.26*** | -.24*** | -.32*** | | |
| 5. Goal Commitment | .05 | .04 | .09 | -.03 | |
| 6. Performance | -.27*** | -.29*** | .15* | .28*** | .29*** |

Note. N = 240; DYB = Do your best goal condition; Easy = Easy goal condition; Difficult = Difficult goal condition; Impossible = Impossible goal condition; * $p < .05$. ** $p < .01$. *** $p < .001$.

standard ANOVA. Results indicated that condition significantly predicted number of uses generated, $F(4, 235) = 16.62$, $p < .001$, partial $\eta^2 = .220$. Pairwise comparisons revealed that the leaderboard condition ($M = 37.52$) significantly outperformed the do-your-best ($M = 23.84$) and easy goal conditions ($M = 22.10$), but performed similarly to the difficult goal ($M = 36.56$) and impossible goal ($M = 41.55$) conditions (see Fig. 1). These results fully support Hypothesis 1.

To test the second hypothesis, hierarchical multiple regression was run with condition and goal commitment in Step 1 and the interaction terms in Step 2. When including a categorical variable in a regression, it must be dummy coded into $g - 1$ code variables, where g is the number of groups (Cohen, Cohen, West, & Aiken, 2003). Therefore, condition was dummy coded into four variables with the leaderboard as the reference group. Each condition variable represents the difference between the leaderboard condition and one other condition with negative numbers indicating that scores are higher in the leaderboard condition. Goal commitment was centered before creating the interaction terms. Each interaction term represents a test of whether the goal commitment slope is different in the two conditions being compared with negative numbers indicating that the slope is stronger in the leaderboard

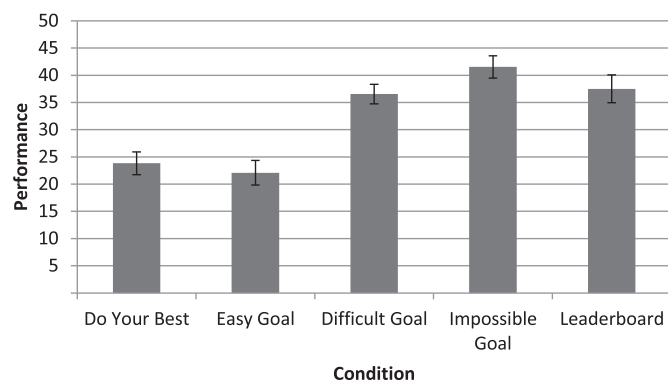


Fig. 1. Graph with standard error bars of the relationship between condition and performance on the brainstorming task.

condition. All assumptions for regression were met, and the results of the regression are presented in Table 3.

Step 1 of the regression represents the test of the difference between conditions controlling for goal commitment. Similar to the results of the ANOVA, the leaderboard condition is significantly different from both the do-your-best ($B = -16.67$, $p < .001$, $sr^2 = .083$) and easy goal conditions ($B = -18.29$, $p < .001$, $sr^2 = .093$), but not different from the difficult ($B = -4.21$, $p = .155$, $sr^2 = .006$) or impossible ($B = 1.93$, $p = .529$, $sr^2 = .001$) goal conditions. In other words, participants in the leaderboard condition generated more uses than participants in the do-your-best and easy goal conditions but generated a similar number of words to participants in the difficult or impossible goal conditions. Consistent with goal-setting theory, goal commitment significantly predicted performance, $B = 7.89$, $p < .001$, $sr^2 = .107$.

Step 2 of the regression represents the test of whether goal commitment functions as a moderator in the way goal-setting theory would predict. Goal-setting theory suggests that goals are only effective if people are committed to them, and performance is maximized when individuals are committed to difficult, specific goals (Locke & Latham, 2002). This suggests that the relationship between goal commitment and performance should be positive in all conditions, but stronger in the leaderboard condition than in the do-your-best or easy goal conditions. The leaderboard vs. do your best contrast approached significance ($B = -8.74$, $p = .087$, $sr^2 = .008$), and the leaderboard vs. easy goal contrast was significant ($B = -11.39$, $p = .024$, $sr^2 = .015$). As shown in Fig. 2, the relationship between goal commitment and performance is strongest in the leaderboard condition. In the do-your-best and easy goal conditions, the relationship between goal commitment and performance is weaker, relative to the leaderboard condition. These results are consistent with Hypothesis 2.

4. Discussion

In this study, we have made three primary contributions. First, we have demonstrated that goal setting can be an effective theoretical framework to explain the success of leaderboards. Decades of research are available on which to base further exploration of leaderboards, and a wealth of additional moderators and mediators are worth exploration. For example, task complexity, which refers to the cognitive load created by pursuit of a goal, is also theorized to moderate the effectiveness of goal setting (Locke & Latham, 2002). Goal setting is generally more effective for simple tasks, with well-defined parameters, in part because it is easier for a person to see the connection between effort and goal achievement (Wood, Mento, & Locke, 1987). Likewise, gamification using leaderboards may be more effective for relatively simple tasks; a leaderboard tracking sales performance is likely to be more effective than a leaderboard tracking managerial success.

Second, we have demonstrated that commitment moderates the success of leaderboards as goal-setting theory would predict (Locke

Table 3

Performance on condition regression results examining the moderating role of goal commitment.

| Variable | Model 1 | | | | Model 2 | | | |
|----------------------------|-----------|------|---------|--------|--------------------|------|---------|--------|
| | B | SE | β | sr^2 | B | SE | β | sr^2 |
| Leaderboard vs. DYB | –16.67*** | 3.10 | –.41 | .083 | –18.71*** | 3.30 | –.46 | .091 |
| Leaderboard vs. Easy | –18.29*** | 3.22 | –.42 | .093 | –20.21*** | 3.41 | –.46 | .100 |
| Leaderboard vs. Difficult | –4.21 | 2.95 | –.12 | .006 | –6.60* | 3.15 | –.18 | .012 |
| Leaderboard vs. Impossible | 1.93 | 3.06 | .05 | .001 | –.27 | 3.27 | –.01 | .000 |
| Goal Commitment | 7.89*** | 1.29 | .33 | .107 | 14.87*** | 4.05 | .63 | .038 |
| DYBxGC | | | | | –8.74 ^a | 5.09 | –.15 | .008 |
| EasyXGC | | | | | –11.39* | 5.00 | –.21 | .015 |
| DifficultXGC | | | | | –4.67 | 4.69 | –.11 | .003 |
| ImpossibleXGC | | | | | –7.98 ^a | 4.82 | –.16 | .008 |
| R^2 | .327*** | | | | .346*** | | | |
| ΔR^2 | | | | | .019*** | | | |

Note. DYB = Do your best goal condition; Easy = Easy goal condition; Difficult = Difficult goal condition; Impossible = Impossible goal condition; GC = Goal Commitment.

^a $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

& Latham, 2002); if people do not believe a leaderboard provides worthwhile goals, leaderboards will not be successful at altering employee behavior. If employees do not believe that the leaderboard is appropriate or worthwhile in the workplace, it is unlikely to affect performance. This can be applied more broadly, as well. If employees do not “buy in” to the idea of gamification, the use of gamification is unlikely to improve work performance.

Third, we have demonstrated that leaderboards are approximately as effective as difficult-to-impossible goals to increase task performance. Individuals faced with a leaderboard are likely to target the top or near-top goals presented on that leaderboard, even without specific instructions to target those goals. It remains unknown if the social component of leaderboards is more motivating than simple goal setting. In this experiment, random initials were chosen to represent each of the prior scorers on the leaderboard. If those initials had represented actual people in the participant's work group, results may have been different. Future research should explore the adequacy of goal-setting theory to explain the success of leaderboards in applied contexts.

4.1. Limitations and directions for future research

This study has three primary limitations. First, it is unclear which specific goals within the leaderboard people actually chose and followed. Although performance ended up at the level of the difficult and impossible goals, this may have been the result of a longitudinal process of goal revision. For example, participants may

have targeted the lowest level and then revised their goal up progressively as their performance improved. Alternatively, participants may have ignored the leaderboard until they had performed to a particular level, only paying attention to it once they were aware of approximately where they would score. Future research should examine the role of goal choice and goal revision longitudinally over the course of a task.

Second, the current study utilized an undergraduate sample performing a short brainstorming task, which may not generalize to other work tasks. Although this paradigm is a common method for studying task performance, it is a relatively simple one, for which there are large effects of goal setting (and as demonstrated here, gamification). As demonstrated by Wood et al. (1987), more complex tasks are likely to produce somewhat smaller effects, although the precise magnitude of this difference is currently unknown. Furthermore, the task was only 12 minutes in length. Participant performance may have been affected by the novelty of the leaderboard in addition to the motivational affordances of goal-setting (Farzan, DiMicco, Millen, Dugan, Geyer & Brownholtz, 2008; Koivisto & Hamari, 2014). Future research should explore a wider variety of tasks with an employee sample, including more complex and longer tasks as commonly seen in modern work environments.

Third, although goal commitment moderates the success of leaderboards, it is unknown how manipulable goal commitment is in the context of gamification. Thus, although it is theoretically meaningful to understand the impact of goal commitment on these relationships, it may not be practically meaningful if goal commitment is less malleable in the context of goal setting than in the context of gamification. For example, because people have pre-existing experience with and exposure to games, goal commitment may be stable before management has the opportunity to affect it. It may be difficult to change the minds of those that believe games are inappropriate for the workplace, in general. To better understand this, future research should identify the antecedents of goal commitment that can be modified by gamification purveyors.

An interesting avenue for future research would be to examine the effect of gamification on the quality of performance. In the present study, participants were instructed to list as many uses as they could rather than being told to thoughtfully brainstorm ideas, but future studies might be designed to examine quality explicitly. The effects of the leaderboard and the moderating role of goal commitment may be stronger when quality is examined, because such an approach requires greater motivation to achieve success. We encourage future researchers to explore tasks in which high quality is rewarded by gamification. For example, participants might be encouraged to identify the most creative solution to a

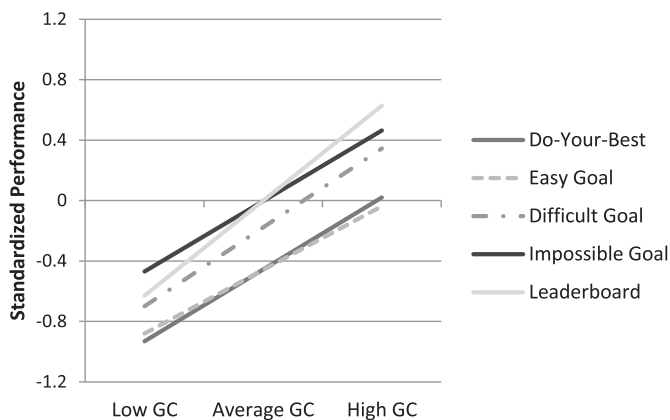


Fig. 2. Graph of the interaction between goal commitment and condition when predicting performance. GC = Goal Commitment. Low goal commitment is 1SD below the mean, whereas high goal commitment is 1SD above the mean.

problem, and originality of the generated solution could be assessed.

4.2. Theoretical and practical recommendations

For researchers, we provide empirical evidence here that goal-setting theory and its sub-theories are valuable frameworks by which to understand the effect of leaderboards on performance. Because explicit and implicit goals are a major component of most gamification interventions, we also contend that goal-setting theory is valuable to understanding the effects of gamification more broadly. Researchers should not re-invent theory when new technologies appear when existing theory already explains their effects. Instead, researchers should identify where those theories are sufficient and insufficient to explain observed effects, augmenting them as appropriate. Those taking an interdisciplinary integrative approach will provide the most value to practitioners seeking to implement these technologies in their own organizations.

For practitioners, we conclude that leaderboards can be an effective approach to gamify task performance. When implemented in support of well-defined tasks using the principles of effective goal setting (SMART goals; Moskowitz & Grant, 2009), task performance is likely to improve. Leaderboards are not the only tools in the gamification toolkit, but they can be effective ones. In general, future research must provide similarly compelling evidence of effectiveness for other types of gamification (Landers, 2015), and it must explore existing theory as a compelling basis for their use to maximize value to researchers and practitioners alike.

Importantly, it should not be concluded from this study that leaderboards are a general solution for the problem of low task performance. Although a powerful tool for directing behavior, goal-setting has significant downsides. By focusing effort narrowly, other desirable outcomes are often deprioritized. For example, Welsh and Ordóñez (2014) found that repeated high performance goals decreased their participants' abilities to self-regulate in other tasks and increased their engagement in unethical behavior. This sort of focus shift may result in systematic harm, including distorted risk management, reduced learning, and decreased intrinsic motivation (Ordóñez, Schweitzer, Galinsky, & Bazerman, 2009). Negative intrinsic motivation effects have already been observed in the gamification context specifically (Mekler, Brühlmann, Opwis, & Tuch, 2013), effectively replicating this goal-setting research. Any gamification intervention, including leaderboards, can be harmful if it directs goals so narrowly that other harms are created (Callan, Bauer, & Landers, 2015). For example, it is easy to imagine a sales office that is so focused upon its salespeople achieving high sales on the office leaderboard that other valid organizational goals, such as ethical sales practices and a positive organizational culture, are ignored. When leaderboards are implemented, they must be targeted to carefully chosen goals, and any potential side effects of that focusing must be monitored continuously.

In closing, we have provided a compelling example of an integrative approach to studying gamification. The use of leaderboards works approximately as well as specific, difficult goal-setting for an idea generation task, and the principles of goal-setting theory should be used to maximize the impact of leaderboards. Goal commitment in particular appears critical to the success of leaderboards. Implementing leaderboards, and likely gamification interventions more broadly, without support from employees is unlikely to lead to increased performance. Managers should integrate the principle of effective goal setting into their gamification, and they should target or at least measure employee goal commitment explicitly before investing substantial resources into gamification interventions, lest their efforts be condemned to failure before they have even begun.

Appendix

Goal Commitment Items – Goal Conditions.

- 1) It's hard to take the goal seriously.
- 2) Quite frankly, I don't care if I achieved the goal or not.
- 3) I was strongly committed to pursuing the goal.
- 4) It wouldn't take much to make me abandon the goal.
- 5) I think this was a good goal to shoot for.

Goal Commitment Items – Leaderboard Condition.

- 1) It's hard to take the leaderboard seriously.
- 2) Quite frankly, I don't care if I scored highly on the leaderboard or not.
- 3) I was strongly committed to scoring highly on the leaderboard.
- 4) It wouldn't have taken much for me to have abandoned getting a high score.
- 5) I think the scores on the leaderboard were good to shoot for.

References

- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139(3), 655–701. <http://dx.doi.org/10.1037/a0029531>.
- 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). Cham, Switzerland: Springer.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.
- Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: practical implications and outcomes. *Computers & Education*, 63, 380–392. <http://dx.doi.org/10.1016/j.compedu.2012.12.020>.
- Eickhoff, C., Harris, C. G., de Vries, A. P., & Srinivasan, P. (2012, August 12–16). Quality through flow and immersion: Gamifying crowdsourced relevance assessments. In *Paper presented at the 35th international ACM SIGIR conference on research and development in information retrieval*, Portland, Oregon, USA.
- Erez, M., & Zidon, I. (1984). Effect of goal acceptance on the relationship of goal difficulty to performance. *Journal of Applied Psychology*, 69(1), 69–78. <http://dx.doi.org/10.1037/0021-9010.69.1.69>.
- Farzan, R., DiMicco, J. M., Brownholtz, D. R. M. B., Geyer, W., & Dugan, C. (2008). When the experiment is over: deploying an incentive system to all the users. In *Paper presented at the symposium on persuasive technology*.
- Farzan, R., DiMicco, J. M., Millen, D. R., Dugan, C., Geyer, W., & Brownholtz, E. A. (2008). Results from deploying a participation incentive mechanism within the enterprise. In *Proceedings of CHI 2008* (pp. 563–572).
- Frayne, C. A., & Geringer, J. M. (2000). Self-management training for improving job performance: a field experiment involving salespeople. *Journal of Applied Psychology*, 85(3), 361–372. <http://dx.doi.org/10.1037/0021-9010.85.3.361>.
- Halan, S., Rossen, B., Cendan, J., & Lok, B. (2010). *High score!-Motivation strategies for user participation in virtual human development intelligent virtual agents* (pp. 482–488). Berlin/Heidelberg: Springer.
- Hamari, J. (2013). Transforming homo economicus into homo ludens: a field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications*, 12(4), 236–245. <http://dx.doi.org/10.1016/j.elerap.2013.01.004>.
- Hamari, J., Koivisto, J., & Pakkanen, T. (2014). Do persuasive technologies persuade? A review of empirical studies. A. Spagnolli, L. Chittaro, & L. Gamberini (Eds.), *Persuasive Technology, LNCS*, 8462, 118–136.
- Hamari, J., Koivisto, J., & Sarsa, H. (January 6–9, 2014). (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii international conference on system sciences*, Hawaii, USA.
- Hinsz, V. B. (2005). The influences of social aspects of competition in goal-setting situations. *Current Psychology*, 24, 258–273.
- Hollenbeck, J. R., Klein, H. J., O'Leary, A. M., & Wright, P. M. (1989). Investigation of the construct validity of a self-report measure of goal commitment. *Journal of Applied Psychology*, 74, 951–956. <http://dx.doi.org/10.1037/0021-9010.74.6.951>.
- Jung, J., Schneider, C., & Valacich, J. (2005). The effects of real-time individual performance feedback and goal setting on computer-mediated group idea generation. *ICIS 2005 Proceedings*, 885–896.

- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: an integrative/apptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74(4), 657–690. <http://dx.doi.org/10.1037/0021-9010.74.4.657>.
- Karoly, P. (1993). Mechanisms of self-regulation: a systems view. *Annual Review of Psychology*, 44, 23–52. <http://dx.doi.org/10.1146/annurev.ps.44.020193.000323>.
- Klein, H. J., Wesson, M. J., Hollenbeck, J. R., & Alge, B. J. (1999). Goal commitment and the goal-setting process: conceptual clarification and empirical synthesis. *Journal of Applied Psychology*, 84, 885–896. <http://dx.doi.org/10.1037/0021-9010.84.6.885>.
- Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, 35, 179–188.
- Koopmans, L., Bernaards, C., Hildebrandt, V., van Buuren, S., van der Beek, A. J., & de Vet, H. C. W. (2012). Development of an individual work performance questionnaire. *International Journal of Productivity and Performance Management*, 62, 6–28.
- Landers, R. N. (2015). Developing a theory of gamified learning: linking serious games and gamification of learning. *Simulation & Gaming*, 45, 752–768.
- 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. 399–423). Surrey, UK: Springer.
- Landers, R. N., & Landers, A. K. (2015). An empirical test of the theory of gamified instructional design: the effect of leaderboards on academic performance. *Simulation & Gaming*, 45, 769–785.
- Latham, G. P., & Baldes, J. J. (1975). The “practical significance” of Locke’s theory of goal setting. *Journal of Applied Psychology*, 60(1), 122–124. <http://dx.doi.org/10.1037/h0076354>.
- Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, 50(2), 212–247. [http://dx.doi.org/10.1016/0749-5978\(91\)90021-k](http://dx.doi.org/10.1016/0749-5978(91)90021-k).
- Lee, F. K., Sheldon, K. M., & Turban, D. B. (2003). Personality and the goal-striving process: the influence of achievement goal patterns, goal level, and mental focus on performance and enjoyment. *Journal of Applied Psychology*, 88(2), 256–265. <http://dx.doi.org/10.1037/0021-9010.88.2.256>.
- Litchfield, R. C. (2008). Brainstorming reconsidered: a goal-based view. *Academy of Management Review*, 33, 649–668.
- Litchfield, R. C., Fan, J., & Brown, V. R. (2011). Directing idea generation using brainstorming with specific novelty goals. *Motivation & Emotion*, 35, 135–143.
- Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior & Human Performance*, 3, 157–189. [http://dx.doi.org/10.1016/0030-5073\(68\)90004-4](http://dx.doi.org/10.1016/0030-5073(68)90004-4).
- Locke, E. A. (1982). Relation of goal level to performance with a short work period and multiple goal levels. *Journal of Applied Psychology*, 67(4), 512–514. <http://dx.doi.org/10.1037/0021-9010.67.4.512>.
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ, US: Prentice-Hall, Inc.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. *American Psychologist*, 57, 705–717. <http://dx.doi.org/10.1037/0003-066x.57.9.705>.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*. New York, NY: Psychology Press.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17, 437–455. <http://dx.doi.org/10.1037/a0028085>.
- Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013). Do points, levels and leaderboards harm intrinsic motivation? an empirical analysis of common gamification elements. In *Proceedings of the first international conference on gameful design, research, and applications (Gamification '13)* (pp. 66–73).
- Moskowitz, G. B., & Grant, H. (2009). *The psychology of goals*. New York: Guilford Press.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: the systematic side effects of overprescribing goal setting. *Academy of Management Perspectives*, 23, 6–16.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.
- Szymanski, K., & Harkins, S. G. (1987). Social loafing and self-evaluation with a social standard. *Journal of Personality and Social Psychology*, 53(5), 891–897. <http://dx.doi.org/10.1037/0022-3514.53.5.891>.
- VandeWalle, D., Brown, S. P., Cron, W. L., & Slocum, J. W., Jr. (1999). The influence of goal orientation and self-regulation tactics on sales performance: a longitudinal field test. *Journal of Applied Psychology*, 84(2), 249–259. <http://dx.doi.org/10.1037/0021-9010.84.2.249>.
- Welsh, D. T., & Ordóñez, L. D. (2014). The dark side of consecutive high performance goals: linking goal setting, depletion, and unethical behavior. *Organizational Behavior and Human Decision Processes*, 123, 79–89.
- White, P. H., Kjelgaard, M. M., & Harkins, S. G. (1995). Testing the contribution of self-evaluation to goal-setting effects. *Journal of Personality and Social Psychology*, 69, 69–79. <http://dx.doi.org/10.1037/0022-3514.69.1.69>.
- Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects: a meta-analysis. *Journal of Applied Psychology*, 72(3), 416–425. <http://dx.doi.org/10.1037/0021-9010.72.3.416>.