Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2019 Proceedings

Human Behavior and IS

Do We Need Different Levels of Badges for Users with Different Participation Levels? A Field Experiment from a Bicycle Commuting Program

Zachary Sheffler
University of Massachusetts Amherst, zsheffler@isenberg.umass.edu

Shawn Curley
University of Minnesota, curley@umn.edu

De Liu
University of Minnesota, deliu@umn.edu

Do We Need Different Levels of Badges for Users with Different Participation Levels? A Field Experiment from a Bicycle Commuting Program

Completed Research Paper

Zachary J. Sheffler
University of Massachusetts
Amherst, MA
zsheffler@isenberg.umass.edu

Shawn P. Curley
University of Minnesota
Minneapolis, MN
curley@umn.edu

De Liu
University of Minnesota
Minneapolis, MN
deliu@umn.edu

Abstract

Using a novel IT infrastructure which measures bicycle ridership within a University, we manipulate the goal condition for the awarding of badges, with "participation" badges for one ride per week, to "challenge" badges for three rides per week. Each condition is targeted at an archetype of rider: those who ride only rarely for the participation badge and thus can benefit from a goal intended to break decision inertia, and those who ride occasionally for the challenge badge who would be challenged by the challenge goal. We find marginal effects of the participation goal among rare riders, with true non-riders being especially difficult to break the state of decision inertia. With infrequent riders, we do not find significant results for the challenge goal but do find a significant increase among infrequent riders presented with the participation goal.

Keywords: gamification, wellness, field experiment, bicycle

Introduction

Gamification, or the application of structures and elements from video games to non-game information systems, has found increased interest among information systems designers and researchers because of the comparatively low costs of gamification compared to other motivational structures. Among the most popular implementations of gamification schemes is the use of badges. Badges have been a staple of gamification research since its earliest conceptions to such a point that badges are one of the prototypical examples of gamification (Liu et al. 2017). Badges have attracted interest from a number of scholars, investigating the effect of badges in contexts from purchase intention (Shang and Lin 2013), to education (Sitra et al. 2017), among others. We contribute to this stream by using a novel information system to study behavior that primarily takes place offline away from an information system, and using online “nudge” characteristics that promote behavior through situational design to motivate positive behavioral changes. The present study uses a field experiment to understand how changes in a badges’ goal conditions affect individuals’ motivations to achieve the goals.

As with other concepts from gamification, badges in non-game information systems are adapted from their use in games. Early uses of badges came in 2005 with the introduction of Microsoft's Xbox 360, which required game developers to incorporate "achievements" into their games that upon completion would contribute a "Gamerscore" amount to a persistent profile shared across games. In this sense, the achievement system constituted a meta-game that proved popular among players, leading to its adoption by competitors, most notably Sony's Playstation 3 and Valve's Steam PC gaming platforms.[1]

Badges began to gain traction in non-game contexts concurrently with the adoption of mobile phones, and were notably used in fitness tracking devices such as the Fitbit, which would award badges for hitting step and stair targets. Community-oriented websites followed suit, offering badges for actions that contributed to the community: Stack Exchange would offer badges for having answers that are voted helpful, for example.

Badges, especially as deployed in non-game contexts, have properties that are of interest to researchers. First is their use as an incentive. It is clear from prior research that badges do have incentive value when deployed within information systems (e.g., Ding, Kim, & Orey, 2017), though the incentive value related to offline activities is less clear. Badges also act similarly to symbolic awards (Kosfeld and Neckermann 2011), which are common motivators that offer no transactional value that can be transferred to another. Instead, badges give recognition as a means of feedback that allows for self-evaluation; and, coincidentally, they are a low-cost option to award.

These motivational properties are of particular interest in contexts where the use of monetary incentives is complicated or rendered impossible by the task or target population. Our study context is one such case, where an entity – the Transportation Office of a University – has a goal to incentivize the increased use of bicycles for transportation to and from campus. Increased bicycle use offers multiple advantages, both to the university and to the riders. Increased bicycle use reduces automotive congestion and parking requirements, improves local air quality, poses fewer risks to pedestrians, and promotes increased health for the riders. However, the nature of the student population makes monetary incentives effectively impossible for legal reasons. We explore the use of badges to motivate bicycle ridership among students at the University.

To accomplish this, we present different ridership goals focused on two distinct archetypes of rider. "Participation" goals are directed at those who do not ride, but with a sufficient motivation, might begin to do so, whom we refer to as "non-riders". The participation goal is based on the minimum measurable activity that can be accomplished: in this case, riding a single day per week. In other contexts, this incentive would be deployed to incentivize the use of a technology or service among those who do not use it but would like to (e.g., people who own a pedometer but are not sufficiently motivated to wear it). Under these circumstances, we expect that (would-be) riders are under the effect of decision inertia (Alós-Ferrer et al. 2016). The participation goal is designed to move the individual to begin riding, in an attempt to overcome the inertia of non-riding, or the identification of oneself as a non-rider (Theodorakis 1994). To differentiate between the differing levels of inertia, we subdivide this group for analysis into those who have no recorded rides and those who have a nonzero amount of recorded rides, but fewer than an average of one per week.

"Challenge" goals are directed at those who do ride sporadically but who, with sufficient incentive, could become more consistent riders, referred to herein as "infrequent riders." The objective here is not one of overcoming inertia or one's self-identification as a non-rider. The objective for these individuals is to enhance their intrinsic motivation to ride, thereby increasing the frequency of their already existing behavior. In doing so, we want to provide a goal that pushes the individual, but is not overly aggressive beyond typical performance levels. For this purpose, the "challenge" goal is calibrated in consultation with the University's transit office to their target of three riding days per week. The three-day goal is designed for occasional, infrequent riders (termed "infrequent riders" in our nomenclature) to bring their activity up to a level that meets the organization's targets.

[1] Each implementation shared similar characteristics broadly, but often used different names, e.g., Playstation Trophies. Throughout this paper we will refer to "badges" generically, following the academic literature.

Each of these two archetypes share qualities which make motivation through badge awards ideal. First, the small marginal benefit that an individual rider obtains from taking an additional ride precludes incentive schemes that have high marginal costs, either in terms of direct benefits or transaction costs. Badges are a low marginal cost option. Second, each population has indicated implicitly a degree of willingness to consider bicycle transportation by joining the program and thus are likely to be receptive to incentives. We inform our incentive design, employing two levels of goals for two types of riders, by investigating the literature on symbolic awards like badges and on previous gamification studies, especially those using badges as a form of symbolic award.

Literature Review

Typical incentive schemes involve compensation which confers an instrumental benefit to the awardee; this is most often cash. Beyond the instrumental benefit, the award of an incentive carries with it – sometimes implicitly – a degree of symbolic value as a result of the award being given. These symbolic features have potential value separate from the instrumental value; we seek to leverage the symbolic value through badge awards that strip away the instrumental value.

Symbolic Awards

Badges are a subset of a larger family of non-economic rewards whose primary purpose is to commemorate an achievement or good performance, termed "symbolic rewards" (Kosfeld and Neckermann 2011). The key distinction between symbolic rewards and incentive structures commonly in use in firms is in the benefits that accrue to the awardee as a component of the award. Symbolic awards confer no instrumental benefit to the awardee, such as monetary incentive or preferential treatment in parking. The benefit is the reward itself and nothing more. This does not necessarily preclude second-order benefits of the award, for example increased status among a peer group; but, badges provide no direct, instrumental, tangible benefit.

Previous studies offer evidence of symbolic awards having a positive effect on the intended outcome variable, such as in reducing absenteeism through awarding good attendance (Markham et al. 2002), and performance in rote tasks such as data entry (Kosfeld and Neckermann 2011). Previous work has also demonstrated positive effects of symbolic awards when offered *ex ante* as a token of commitment (Baca-Motes et al. 2012). As such, there is evidence from prior research that rewards that share critical properties with badges motivate agents, even when it is understood that there is no instrumental value to be gained from earning the badge. However, badges do have unique properties that are not addressed in the previous research into symbolic rewards, since they are online awards that are digital in nature. So, we turn our attention to badges themselves as a gamification mechanism and their motivating elements.

Gamification and Badges

The use of badges has attracted considerable interest from researchers. Badges share the properties of symbolic awards as discussed above: They provide no direct benefit to the awardee other than the badge itself. No necessary criteria exist for designing and awarding badges, though within the literature some common properties emerge. Among these are explicit and stated goal conditions to receive the badge and a computerized image. Furthermore, badges are exclusively framed as positive rewards; current implementations do not use badges as negative incentives when a person performs undesirable behavior.

Other badge features are commonly, but not exclusively, used in gamification studies. Among the most common are public displays of badges as a component of a user's profile (Hamari 2017), use of accompanying or "flavor text" – a short description of the badge or accomplishment – along with the badge (Ding et al. 2017), and social metrics indicating the success rate for other users (Denny 2013). While the specific features of badges as implemented in video games and major gamification systems are diverse, the important features remain fixed, and we focus our attention on these features. Specifically, badges operationalized in this study have an image, a notification of award, and a "win" goal condition which is known, *ex ante*, to the agent. Within these constraints, the manipulation of the goal condition presents opportunities to motivate individuals which derive from phenomena rooted in the choice architecture and motivational psychology literatures, to which we now turn our attention.

Goals

For the purposes of this study, we draw on behavioral literature for guidance on how badges might motivate individuals, particularly for promoting wellness activities where our study is rooted. Our interest is in the role that differing goals attached to badges have in motivating two types of users—those who are active but want to be more so and those who are mostly inactive and want to begin. Specifically, we review literature related to goals and decision making, as well as "decision inertia", or the tendency to repeat previous choices, of being stuck in an existing pattern of behavior (or non-behavior).

Goals play an important role in decision making. Explicit, challenging goals enhance a person's motivation compared to non-specific goals such as "do your best" (Locke and Latham 2002). Goals also act as reference points against which performance can be judged (Heath et al. 1999), which increase the likelihood of goal attainment. Specific goals further act as feedback mechanisms, which increase feelings of mastery on the part of the person, which will itself positively affect motivation (Ryan and Deci 2000). As such, specificity of a goal is important in motivating behavior.

Beyond the general properties of setting a goal, the nature of the goal condition itself will affect behavior. Goals which require minimal further effort to be attained are termed *proximal*, while goals which still require significant effort to be attained are termed as *distal*. The proximity of a goal will affect the person's motivation with respect to that goal: the goal-gradient hypothesis finds that motivation increases as a goal becomes more proximal. Initial studies focused on rats and maze learning (Hull 1932), though more recent research has applied this to human consumers, showing for example an increased interest in purchasing coffee when closer to completing a loyalty-program punch card (Kivetz et al. 2006).

These theories of goal motivation have been applied to badges with promising results. Of particular interest to the present study is Mutter & Kundisch's (2014) application of the goal-gradient hypothesis to online badges in the context of user behavior on an online question-and-answer site before and after an exogenous change to the thresholds to earn successive levels of badges. The nature of the change effectively randomized users' progress toward the next badge level, and subsequent behavior observed on the site was consistent with expectations given by the goal-gradient hypothesis, namely that users who were closer to the end of a goal path exhibited greater levels of posting behavior relative to users who were closer to the midpoint. The general lesson is that a goal is more effective when it is tailored to the present state of the individual being motivated. However, as with most badge studies, the goal of the program is to increase engagement on the site itself rather than change real-world, offline behavior, and this remains a gap in the literature.

For non-riders, an additional motivational factor can be operative. When considering participants who do not exhibit any riding behavior, the concept of decision inertia can apply. Decision inertia is the tendency to repeat previous choices, even when these choices are suboptimal (Alós-Ferrer et al. 2016). The theoretical basis comes from dual-process cognition, wherein the automatic system uses previous decisions as a default (Strack and Deutsch 2004), as well as from a general psychological preference for consistency. However, the majority of studies in these areas have been focused on probability tasks: Alós-Ferrer et al. (2016) focused on the role of Bayesian updating on ball draws, for example. A related idea arises from the role of self-identity upon decision making, primarily studied in a consumer marketing setting (Reed et al. 2012 provide a review). A self-identity is a category with which an individual self-identifies and that is then used to monitor their own behavior. Marketers apply this idea in trying to get consumers to self-identify with their brands to create positive associations with their products. To the extent that participants self-identify as non-riders, this would promote decision inertia and serve as a barrier to action. Of closest relevance to the current work is the finding of Theodorakis (1994) that exercise behavior can constitute a form of self-identity that relates to actual exercise behavior. His study was survey-based and predictive, however, and could not address possible causality nor was it interested in testing behavioral change. We contribute to the literature by considering behavioral interventions with the goal of supporting wellness behavior within an experimental field setting.

Hypotheses

We consider two types of goals, each targeted at one of the two archetypes discussed earlier. A participation goal is designed primarily for "non-riders" and a challenge goal is designed primarily for "infrequent riders."

We explicitly consider in our hypotheses how each type of user would react to these two goal types, singly and together.

Among non-riders, the effects of decision inertia are predominant. Motivating a first ride can be difficult. Extra motivation is needed to provide the impetus to exercise once compared to the motivation needed to exercise for a second or later instance. Providing only a challenge goal is not expected to provide sufficient motivation. For the non-user, the challenge goal is a more distal goal for which motivation is low based on the goal-gradient hypothesis. In contrast, for the non-rider a participation goal is an achievable proximal goal. Thus, based on the goal-gradient hypothesis that the proximity to the goal matters, we hypothesize that for a significant number of non-riders the motivation of the proximal participation goal will be more likely to provide an incentive to overcome decision inertia and ride.

More generally, the idea is that a low-effort badge leverages the goal-gradient hypothesis to motivate users who are in an inertial state – in this case, not exercising – to break their decision inertia. We combine this with an expectation that for these users, the challenge badge will be perceived as requiring too much effort to offer sufficient motivation. These expectations lead to the following hypotheses for non-riders:

> H1: Non-riders who can earn a participation badge will increase their ridership compared to those with no badge award available.

> H2: Non-riders who can earn a challenge badge will not increase their ridership compared to those with no badge award available.

As part of the factorial experimental design a third case was investigated: that of providing both participation and challenge badges together. This condition is important both theoretically and practically. From a practical perspective, it allows us to determine whether both goals just can be used simultaneously without regard for the participant, or whether it is necessary to tailor the incentive to the individual for a positive effect. Indeed, when both participation and challenge badges are available to non-riders, the expectation is somewhat more speculative. If the effect is just summative of the individual effects, then the participation badge will have the usual effect and the challenge badge little effect, leading to the expectation that the use of both badges is roughly equivalent to having the participation badge alone. An alternative is that having both badges available creates a comparison effect leading to a contrast between the awards in their perception. There is ample evidence across a variety of situations that perception and judgment are comparative processes, for example, contrast effects in visual perception (Plous 1993) and decoy effects in judgment (Highhouse 1996; Scarpi and Pizzi 2013). It is possible that adding the challenge goal creates a comparison against which the lesser participation goal becomes contrasted. This process conceivably could devalue the participation goal, lessening its motivational effect and providing insufficient incentive to counter decision inertia. In this case, the effect of having both badges available would be less than that of having a participation badge alone available. Thus, the effect of offering both badges is of general theoretical interest, as informing the motivational implications of these possibly competing motivations.

From a practical standpoint, perhaps the best case scenario is that offering both badges will be equivalent to the offer of a lower-tier badge alone. This would allow a practice of offering badges without having to consider the receiver, an operationally easier implementation. The alternative is that offering both badges reduces or completely eliminates their effectiveness, arguing for a more nuanced, tailored implementation. We posit these as two contrasting hypotheses for the joint effect, the results for which will also help to inform badge implementation for information systems designers.

> H3a: Non-riders who can earn a participation and a challenge badge will have a ridership equivalent to those who only can earn a participation badge.

> H3b: Non-riders who can earn a participation and a challenge badge will have a ridership less than those who only can earn a participation badge.

Infrequent riders, who already average at least one ride per week (but no more than three) in the beginning period, when presented with a participation goal, are unlikely to be affected by the addition of an incentive for behaviors which they already perform. The inertia for the infrequent rider is toward riding some during the week, so no initial inertia of inactivity needs to be overcome. On the other hand, a challenge goal can provide an incentive benefit as the challenge goal becomes more proximal with each ride, based on the goal-gradient hypothesis. Hence, we expect the effects of the badges to be opposite of those for non-riders as formulated by the following hypotheses for the effects of each badge singly:

H4: Infrequent riders who can earn a participation badge will not increase their ridership compared to those with no badge award available.

H5: Infrequent riders who can earn a challenge badge will increase their ridership compared to those with no badge award available.

Considering infrequent riders with both badges available, we reiterate that we hypothesize no effect for the participation badge as the behavior being rewarded is already being performed, nor does a non-riding state of inertia exist for which the badge acts as a nudge to break that loop. Further, unlike non-riders, we expect that infrequent riders will not be adversely affected by the offering of two badges given that before the badge implementation, infrequent riders would have, on average, earned the lesser badge. Thus, there is no comparable disincentive that having both would be expected to provide. We largely expect this badge to be ignored by the infrequent rider. So, unlike the non-rider, we do not expect the presence of a participation badge to have any crossover effect on the effect of the challenge badge. We hypothesize that the challenge badge's main effect will remain, leading to the following hypothesis:

H6: Infrequent riders who can earn both a challenge and a participation badge will have a ridership equivalent to those with only a challenge badge award available.

The final group, frequent riders, we do not target with any specific treatment. Since this group already averages at least three rides per week, no intervention is needed to increase ridership, and it is likely that riders at this level are motivated by factors orthogonal to an external reward system, such as intrinsic interest or lack of other transportation options. As such, we do not expect the badges to have any effect on this group and we exclude this group from our hypotheses, though their analyses are included for completeness and contrast.

Methods

Context and Description of Technology

The university, in conjunction with a local bike rack manufacturer, developed technology for monitoring bicycle commuting, which will be referred to as the Bicycle Commuting (BC) program, with a goal of providing measurement for reimbursement of insurance premiums for insurance policyholders who commute frequently by bicycle. This is a desirable behavior from the perspective of the university, which offers other incentives for physical activity, such as gym attendance. Furthermore, this is desirable from the perspective of the participants, who opt into the program and are thus nominally interested in increasing their ridership behavior.

After developing this technology, the program expanded to include students, who do not participate in the university's employee insurance plan.[2] While there is no direct financial incentive, the University has deemed it to be in the interest of the community to increase bicycle ridership among students to promote healthy behavior and decrease automobile congestion on campus and the surrounding area. However, because of tax implications, the university cannot issue direct cash payments to students to incentivize behavior. As a workaround, students who meet a minimum number of riding days (twelve per month, or approximately 3 days per week) are entered into a monthly lottery for a $10 gift card, donated by local businesses. Students, for their part, reap the benefits of increased physical activity, including increases in affect, decreases in weight, and higher academic performance (Stubbe et al. 2007).

The BC equipment has two components: first is a radio-frequency identification (RFID) chip which attaches to bicycle wheels at the spokes. These are relatively low-cost – approximately $3 per chip – and are distributed for free at tabling events and through the Parking and Transportation office to students, staff, and faculty who choose to participate. This program is strictly opt-in; there are no negative consequences for not participating in the program. Each RFID chip has a unique serial number that is associated with a profile within the information system.

[2] Students may optionally enroll in a university-administered plan which is separate from the plan used for staff and faculty.

The second component are "readers" which are placed in high-traffic areas, typically entry and exit points to areas with high bicycle traffic, such as major entry and exit points to college campuses and downtown areas. These readers have directional antennae that detect the RFID chips and record the serial numbers and the time of the scan. A wireless modem on the reader pole uploads these data, combined with a unique ID from the reader, to a central database.

Participants

The participants were students at a large University in the Midwest who had opted into the BC program and had installed a RFID tag on their bicycle. This was further narrowed to participants who had recorded at least one ride (hereafter referred to as a scan) during the academic year to avoid including former students who moved away from the University area or had chosen to opt-out of the program by removing their RFID tag. Users who chose to opt-out from receiving the badge email used in the treatment groups of the study were excluded from analysis. Following these actions, a total of 2,616 participants were identified for the study.

Procedure

Participants were randomly sorted into one of five groups in a 2 (one-ride participation badge available: yes/no) x 2 (three-ride challenge badge available: yes/no) + 1 (control: no encouragement) design. Users were randomized into conditions as indicated by Table 1. Data are balanced with the exception of the conditions wherein both badges are unavailable to the user. The participants in the two conditions without badges were roughly equally split between the two control condition variants.

Table 1: Distribution of Participants across Conditions

	Participation Goal	No Participation Goal
Challenge Goal	675	640
No Challenge Goal	663	311
Control (info only)	327	

The one-ride participation goal incentive was presented in the form of a gold badge that was shown, or indicated as unearned by the image of a closed lock. The three-ride challenge goal incentive was presented as a diamond badge, or by a closed lock if unearned. The no challenge/no participation badge condition and the fifth "no encouragement" condition are used as controls; for these, no badge was indicated as being available and no badge or lock images appeared. In the "no/no" control condition, no badges could be earned, but the users were still encouraged to ride with the following text on the display: "Try to ride at least once next week! Three or more times is better!" The "no encouragement" condition did not contain this text, only showing the number of rides in the previous week.

Riders could access their badge awards in two ways. The intended primary route was through e-mail: each Sunday, a personalized email was delivered to the registered email address displaying the rider's name and number of riding days over the previous week. Additionally, for treatment groups receiving a badge, the email contained the badge image and text of the highest badge earned, or for users who did not earn a badge, a lock icon indicating that a badge was there but hidden. All users saw an encouraging message which varied according to which badges were available, with the exception of the no badge-info only group, which saw no encouraging information with the aim of separating out the effect of the encouraging messages.

The second route available to riders was a web dashboard on the BC website, which all participants needed to visit at least one time to register their RFID chip. The dashboard could be accessed at any time. Aside from minor changes in color and font to comply with email security standards, the presentation of the email and the relevant pane on the website were identical. The website also contained a link to share their badge to Facebook (which was logged when clicked, though the link was deactivated during the study to avoid leakage to any social-comparison effects), and another link to opt-out of the emails. Badges were available only to the user and could not be shared. This was to minimize the possibility of contamination of treatments. [Images are available but removed to blind for peer review]

Analysis

Data were filtered to a seven-week period during the Fall semester immediately preceding the Thanksgiving break (not including the short week which includes Thanksgiving). Because the shifts in regional weather patterns roughly coincide with Thanksgiving and Spring Breaks, students often take their bicycles home at Thanksgiving and retrieve them after Spring Break. The eight-week window was chosen as it coincides with the remaining instructional time between Spring Break and the end of the Spring term.

Since we hypothesize effects based on the archetype of rider, we rely on a pre-manipulation measurement period where we observe riding behavior spanning four weeks at the beginning of the semester to determine rider archetype. Riders who averaged less than one ride per week over that period were categorized as "non-riders". To account for the inertial differences between riders who have not ridden at all during the observation period, and those who have demonstrated any proclivity to ride, "non-riders" were further subdivided into two groups, "zero-riders" and "rare" riders, with the former having ridden zero times during the study period, and the latter having less than one average weekly ride but greater than zero. For the purposes of our hypothesis development, we continue to consider both groups "non-riders." Riders whose average was greater than or equal to one, but less than three rides per week were categorized as "infrequent riders", and those with three or more rides per week were categorized as "frequent riders", whom we did not target with manipulation. Using these cutoff points, participants were sorted into categories as indicated by Table 2.

Table 2: Distribution of Rider Archetypes

	Zero-Riders	Rare Riders	Infrequent Riders	Frequent Riders
Count	1261	298	457	600
Mean Rides/Week (pre-period)	0.000	0.448	1.855	4.227
Mean Rides/Week (study)	0.132	0.387	1.275	3.180

Panels were created with (anonymized) user IDs, week IDs, the total number of scans, and the number of unique days a scan occurred. We focus on days a scan occurred as the best measure given the nature of the implementation of the RFID technology. The placement of BC poles is concentrated at "choke points" where people are likely to ride a bicycle: major entrances to the campus, for example. It is a strong assumption, and likely a poor one, that a larger number of scans would imply more riding behavior. A person riding one mile across campus may record three scans, while a person riding 20 miles to the edge of campus may only record one. Furthermore, the goal of the study is to increase ridership generally. Rewarding riders for recording scans has the potential to change riding behavior to maximize the number of scans, rather than change the targeted behavior of increasing ridership. For these reasons, manipulations and communications with users are presented in terms of days riding.

Each hypothesis was framed in terms of specific archetypes, which we operationalize in our model:

$$\begin{aligned} days\ riding_{archetype} &= \beta_{archetype,0} + \beta_1 participation_{archetype} + \beta_2 challenge_{archetype} \\ &+ \beta_3 \left(participation_{archetype} * challenge_{archetype}\right) + \epsilon_{archetype} \end{aligned}$$

The input data for the model includes all behavior over the seven-week period arranged in a cross-section. The variables *participation* and *challenge* are 0/1 indicator variables taking the value 1 for those in the conditions receiving the respective badge. The interaction term captures the possible, additional joint effect of having both badges available together. Finally, the model includes the standard error term. Because this is a Poisson model, we avoid the use of fixed effects at the person-level to avoid the incidental parameter problem. Further, we use category-wise regressions per archetype, as recommended in Holgersson, Nordström, & Öner (2014) to avoid interpretive issues with three-way interactions from using category-wise 0/1 variables. We evaluate model fit with Cox and Snell's Pseudo-R^2, which compares the difference between a null model and the fitted model (Haigh et al. 1990).

Results

Riders were binned according to their ridership behavior in the four weeks preceding the study period. Averages per group are shown in Table 2. The two control conditions--no encouragement and the no-participation, no-challenge goal conditions— did not significantly differ ($t = -0.406, p = 0.685$), as such we collapse the two conditions into a single no-goals control condition designated in the model as the base case where *participation* = 0 and *challenge* = 0. The ensuing model fit is summarized in Table 3. As indicated above, the model is fit separately for each archetype. Weeks in Table 3 utilize Week 1 as a baseline and as such coefficients should be interpreted as change relative to behavior during Week 1. Although not the subject of any hypothesis, the analysis for the frequent rider archetype group is included for comparison. As indicated in Table 3 and as expected, none of the badge manipulations had an effect for this group.

Table 3: Regression Results, Coefficients with (standard errors)

DV: Rides per Week	Zero-Riders	Rare Riders	Infrequent Riders	Frequent Riders
Intercept	-1.849 (0.051)***	-1.110 (0.076)***	0.191 (0.034)***	1.148 (0.017)***
Participation (1 = yes)	-0.228 (0.076)**	0.180 (0.106)†	0.118 (0.045) **	0.033 (0.025)
Challenge (1 = yes)	-0.188 (0.078)*	0.227 (0.102)*	-0.040 (0.480)	0.001 (0.024)
Participation x Challenge	-0.117 (0.121)	-0.186 (0.142)	0.024 (0.063)	-0.028 (0.035)
N	1261	298	457	600
Cox-Snell Pseudo-R^2	.0044	.0031	.0056	.0005

† p < 0.1,* p < 0.05. ** p < 0.01. ***p < 0.001.

Beginning with the non-rider archetype, these riders were affected by the participation badge, but not fully as hypothesized. Hypothesis H1 predicted that the participation badge would increase the riding behavior of non-riders compared to those outside of the condition. The participation badge in fact caused a significant decrease in ridership ($p = 0.002$) among zero-riders but a marginally significant *increase* among rare riders ($p = 0.089$). As such, we consider H1 to not be supported. Hypothesis H2, which stated that the challenge badge would not affect ridership for non-riders, similarly is directionally opposite and significant for zero-riders ($p = 0.017$), but significantly increases the ridership of rare riders ($p = 0.026$), and as such we consider H2 to be partially supported. H3a and H3b offered contrasting possibilities, namely that being offered both badges would have no effect over the participation badge (H3a) or a negative effect relative to the participation badge (H3b). No significant interaction between the participation and challenge badges were present for either zero-riders ($p = 0.332$) or rare riders ($p = 0.190$) consistent with hypothesis H3a and not supportive of H3b. As such, the null result hypothesized as H3a could not be rejected.

For the infrequent rider archetype, the results were not fully consistent with the hypotheses either. Infrequent riders were hypothesized to not be affected by the participation badge (H4); but, they did in fact exhibit significantly increased ridership when offered a chance for a participation badge ($p = 0.005$). Hypothesis H5 was also not supported: The challenge badge did not increase ridership as hypothesized by H5 ($p = 0.507$). As with the non-riders and consistent with hypothesis H6, no interaction effect was found between the participation and challenge badges ($p = 0.673$).

Discussion

Our goal in this study was to account for the heterogeneity that exists in the real world by offering different levels of badges that are specifically targeted to different populations. While we utilize the context of bicycle ridership, levels were chosen in a generalizable manner: participation awards for performing the minimum observable behavior afforded by the technology, and challenge awards for performing that same behavior at a level determined by administrators to be the desirable behavior for their target population. We observe, but do not attempt to specifically motivate, riders who meet or exceed the desired threshold and note that no evidence of a demotivational effect is present in our data.

We offered hypotheses based on the relative effect of the badge offers on different archetypes of riders, with each badge targeting a specific archetype. With respect to the non-rider archetype, which averaged fewer than one ride per week in the first four weeks of the semester, we posited that the participation badge would increase ridership (H1). Splitting non-riders into two groups, we find opposite results, with zero-riders having a significant decrease in ridership ($p = 0.003$) and rare riders having a marginally significant increase ($p = 0.089$). The challenge badge showed a similar pattern: negative for zero-riders ($p = 0.017$), and positive for rare riders ($p = 0.026$) despite the hypothesis of no effect (H2). As such, these findings are clearly worthy of further study. The effect runs counter to the expectation that a participation badge would help break the decision inertia of being a non-rider and promote an onset of riding behavior. Not only did this not occur, but the indication was that offering a participation badge was a disincentive to riding among zero-riders. With the rare riders, who would still be subject to some level of decision inertia but had a demonstrated proclivity to ride, these badges did appear to have efficacy. We interpret this as evidence of the difficulty of overcoming decision inertia and note that a stronger incentive may be needed to induce zero-riders to ride. Indeed, the participation badge may just provide a reminder of their lack of participation, serving as a significant disincentive to start participating.

Interestingly, the offering of a participation badge led to an opposite effect for infrequent riders, again counter to the expectations based on existing theory. For infrequent riders, decision inertia toward not riding was not expected to be a factor. Yet, a participation badge for these riders that are already riding, increased ridership. One possibility is that receiving a badge for behavior they were already engaged in doing provided a recognition of that activity as a form of positive feedback, leading to an increase in ridership. This is a possibility that would be consistent with extant goal-setting literature (Locke 1968). In contrast, offering a challenge badge, which constitutes a target to strive toward, did not increase ridership. This set of findings raises an issue for further investigation regarding badges as incentives. A more direct study targeted at contrasting the badge as offering a target goal vs. the badge as a form of positive feedback is suggested.

From a practical standpoint, the findings clearly indicate that badges offered to the full population will have varying effects depending on the rider. There is no single type of award that is indicated at being universally effective. Our research highlights this for practitioners: the heterogeneity of the target populations requires careful consideration of their idiosyncrasies with respect to the goals of the treatment.

We also note that badges cannot be assumed to be monotonically beneficial, and that their deployment may come with adverse effects. While the negative effect in this case was only marginally significant, it remains an important consideration that badges may in fact have de-motivational effects. In this regard, although not a direct subject of our investigation, the lack of an effect on frequent riders carries with it an important implication for practitioners: The addition of badges had no significant effect on those who are already performing beyond the level which is being targeted. This finding alleviates one concern for badge design: The possibility that adverse consequences will present themselves among non-targeted populations. This seems not to be the case among this group, who are already performing at a sufficiently high level and do not show any significant change in behavior with the introduction of a badge.

Implications for the populations which are targeted present different challenges. Zero-riders did not respond to the incentive. The introduction of a badge was not a sufficient motivator to overcome the inertia of never riding to performing any riding behavior. Rare riders already demonstrated some proclivity to ride and might respond to the participation incentive showing an effect with marginal significance. This suggests that offering badges for minimal behavior can be effective in cases where the recipient has already demonstrated some minimal level of the target behavior, but the motivational power of a badge may not be strong enough to overcome the comparatively more powerful inertia of going from zero to nonzero.

Among infrequent riders, the addition of a participation badge did appear to reinforce existing behavior, though challenge badges did not increase behavior. This indicates that the efficacy of badges is likely to be reliant on the amount of behavior change which is required to earn them. This is consistent with the lack of effect on zero-riders of the participation badge but evidence of some effect on rare riders.

This presents implications for future research as well. First, we contribute to the small but expanding literature which studies gamification and badges in randomized field experiments. Second, we explore the limits of the motivational power of badges and find evidence that there is a maximum amount of behavior change which a badge can induce, and one which the present study may have exceeded. At the same time, it is clear that badges do have some motivational power, especially among a subset of users. The boundaries of the behavioral change which badges can induce is still an open question and one worth exploring.

References

Alós-Ferrer, C., Hügelschäfer, S., and Li, J. 2016. "Inertia and Decision Making," *Frontiers in Psychology* (7:169), pp. 1–9. (https://doi.org/10.3389/fpsyg.2016.00169).

Baca-Motes, K., Brown, A., Gneezy, A., Keenan, E. a., and Nelson, L. D. 2012. "Commitment and Behavior Change: Evidence from the Field," *Journal of Consumer Research* (39:5), pp. 000–000. (https://doi.org/10.1086/667226).

Denny, P. 2013. "The Effect of Virtual Achievements on Student Engagement," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, CHI '13, New York, New York, USA: ACM Press, p. 763. (https://doi.org/10.1145/2470654.2470763).

Ding, L., Kim, C., and Orey, M. 2017. "Studies of Student Engagement in Gamified Online Discussions," *Computers & Education* (115), pp. 126–142. (https://doi.org/https://doi.org/10.1016/j.compedu.2017.06.016).

Haigh, J., Cox, D. R., and Snell, E. J. 1990. "Analysis of Binary Data," *The Mathematical Gazette*. (https://doi.org/10.2307/3618895).

Hamari, J. 2017. "Do Badges Increase User Activity? A Field Experiment on the Effects of Gamification," *Computers in Human Behavior* (71), pp. 469–478. (https://doi.org/https://doi.org/10.1016/j.chb.2015.03.036).

Heath, C., Larrick, R. P., and Wu, G. 1999. "Goals as Reference Points.," *Cognitive Psychology* (38:1), pp. 79–109. (https://doi.org/10.1006/cogp.1998.0708).

Highhouse, S. 1996. "Context-Dependent Selection: The Effects of Decoy and Phantom Job Candidates," *Organizational Behavior and Human Decision Processes* (65:1), Academic Press, pp. 68–76. (https://doi.org/10.1006/OBHD.1996.0006).

Holgersson, H. E. T., Nordström, L., and Öner, Ö. 2014. "Dummy Variables vs. Category-Wise Models," *Journal of Applied Statistics* (41:2), pp. 233–241. (https://doi.org/10.1080/02664763.2013.838665).

Hull, C. L. 1932. "The Goal-Gradient Hypothesis and Maze Learning.," *Psychological Review* (39:1), pp. 25–43. (https://doi.org/10.1037/h0072640).

Kivetz, R., Urminsky, O., and Zheng, Y. 2006. "The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention," *Journal of Marketing Research* (43:1), pp. 39–58. (https://doi.org/10.1509/jmkr.43.1.39).

Kosfeld, M., and Neckermann, S. 2011. "Getting More Work for Nothing? Symbolic Awards and Worker Performance," *American Economic Journal: Microeconomics* (3:3), pp. 86–99. (https://doi.org/10.1257/mic.3.3.86).

Liu, D., Santhanam, R., and Webster, J. 2017. "Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems," *MIS Quarterly* (41:4), MIS Quarterly, pp. 1011-1-A4. (http://login.ezproxy.lib.umn.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=buh&AN=125996703&site=ehost-live).

Locke, E. A. 1968. "Toward a Theory of Task Motivation and Incentives," *Organizational Behavior and Human Performance* (3:2), pp. 157–189. (https://doi.org/10.1016/0030-5073(68)90004-4).

Locke, E. A., and Latham, G. P. 2002. "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey.," *American Psychologist* (57:9), pp. 705–717. (https://doi.org/10.1037/0003-066X.57.9.705).

Markham, S. E., Scott, K. D., and McKee, G. H. 2002. "Recognizing Good Attendance: A Longitudinal, Quasi-Experimental Field Study," *Personnel Psychology* (55:3), pp. 639–661.

Mutter, T., and Kundisch, D. 2014. "Behavioral Mechanisms Prompted by Badges: The Goal-Gradient Hypothesis," in *ICIS 2014 Proceedings*. (http://aisel.aisnet.org/icis2014/proceedings/SocialMedia/12).

Plous, S. 1993. *The Psychology of Judgment and Decision Making*, New York, NY. (https://psycnet.apa.org/record/1993-97429-000).

Reed, A., Forehand, M. R., Puntoni, S., and Warlop, L. 2012. "Identity-Based Consumer Behavior," *International Journal of Research in Marketing*. (https://doi.org/10.1016/j.ijresmar.2012.08.002).

Ryan, R. M., and Deci, E. L. 2000. "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions.," *Contemporary Educational Psychology* (25), pp. 54–67. (https://doi.org/10.1006/ceps.1999.1020).

Scarpi, D., and Pizzi, G. 2013. "The Impact of Phantom Decoys on Choices and Perceptions," *Journal of Behavioral Decision Making* (26), pp. 451–461. (https://doi.org/10.1002/bdm.1778).

Shang, S. S. C., and Lin, K. Y. 2013. "An Understanding of the Impact of Gamification on Purchase Intentions," in *Proceedings of the Nineteenth Americas Conference on Information Systems*, pp. 1–11. (http://aisel.aisnet.org/amcis2013/eBusinessIntelligence/RoundTablePresentations/11).

Sitra, O., Katsigiannakis, V., Karagiannidis, C., and Mavropoulou, S. 2017. "The Effect of Badges on the Engagement of Students with Special Educational Needs: A Case Study," *Education and Information Technologies* (22:6), pp. 3037–3046. (https://doi.org/10.1007/s10639-016-9550-5).

Strack, F., and Deutsch, R. 2004. "Reflective and Impulsive Determinants of Social Behavior," *Personality and Social Psychology Review* (8:3), pp. 220–247. (https://doi.org/10.1207/s15327957pspr0803).

Stubbe, J. H., de Moor, M. H. M., Boomsma, D. I., and de Geus, E. J. C. 2007. "The Association between Exercise Participation and Well-Being: A Co-Twin Study," *Preventive Medicine* (44:2), pp. 148–152. (https://doi.org/10.1016/j.ypmed.2006.09.002).

Theodorakis, Y. 1994. "Planned Behavior, Attitude Strength, Role Identity, and the Prediction of Exercise Behavior," *The Sport Psychologist* (8:2), pp. 149–165. (https://doi.org/10.1123/tsp.8.2.149).