Gamification Strategies of Badges as a Solution to Issues with Cash Incentives

A Dissertation Submitted to the Faculty of the University of Minnesota by

Zachary J. Sheffler

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Shawn P. Curley, Advisor

August 2018

Acknowledgements

First and foremost, I owe an enormous debt of gratitude to my advisor, Shawn Curley, whose patience in refining the contents of this document proved to be seemingly boundless. His guidance and mentorship over the past five years has been invaluable.

Nearly as critical to the completion of this was De Liu, who played an instrumental role in developing the theoretical bases for each of the experiments and provided a great deal of technical expertise in realizing the experimental designs.

Mani Subramani, committee member and sometime collaborator, similarly played an important role in my development as a scholar. Additionally, my sincere thanks to Michael Harwell for taking the time out to serve on my committee and providing valuable feedback at each of the stages of completion of this dissertation.

I've had the good fortune to be housed within the Minnesota Information and Decision Sciences Department with eminent scholars who have pushed my thinking without fail; Ching Ren, Ravi Bapna, Gord Burtch, Alok Gupta, Gedas Adomavicius, and Brad Greenwood all (in no particular order) had a significant impact on my development during the doctoral program.

Among the fellow students, colleagues, and friends of mine who are too numerous to thank, I'd single out Mochen Yang, JaeHwuen Jung, Dustin Abbott, Jessica Freeman, Lauren Mitchell, and Jonathan Borowsky, each of whom helped me keep perspective during the long process of the doctoral program.

I want to make sure to credit the patience of my wife, Jane Sheffler, who dutifully put up with the various unpleasantries that were imposed on her as a result of my status as a doctoral student, for which she has been handsomely rewarded by an acknowledgment in a dissertation, which I hope is sufficient.

Finally, to James, my son, I want to express my deep gratitude for making me realize that being well-rested is over-rated. I dedicate this work to him.

For James
"What's That?"

Contents

Table of Tables

Table of Figures

Chapter 1 - Introduction

The use of cash as a motivator has long been the standard instrument to motivate people to perform actions that they would not otherwise do on their own. Principal-agent theory is framed explicitly in terms of cash (Eisenhardt, 1989), and cash is typically used synonymously with extrinsic motivation (Benabou & Tirole, 2003). Experimentalists nearly universally use cash as an additional motivator across several disciplines, including psychology (Goette & Stutzer, 2008) and political science (Imai & Goldstein, 2007), among many others.

It is true that cash makes for an effective motivator given its place in modern society and its obvious value, but these positive features also cause some drawbacks. One problem observed with cash incentives is motivation crowding, where the introduction of an external incentive supplants a person's internal motivation (Frey & Jegen, 2001). This leads a dependence relation on the extrinsic motivation and a resulting drop in motivation, after the incentive is removed. Illustrating this, in Charness & Gneezy's (2009) study of monetary incentives and gym attendance, while the cash incentive increased the gym attendance of those least likely to go to the gym, even after the removal of the incentive; after the study period where the incentive was removed, those who regularly attended experienced a drop in their attendance. This leads to incentive designers either unknowingly demotivating a subset of their population, knowingly making a tradeoff between the high and low performers or committing to cash incentive in perpetuity when implementing a cash incentive scheme.

A second issue with the focus on monetary incentives is that there are times when issuing a cash incentive becomes infeasible because of the high marginal transaction costs being out of proportion to the level of the incentive. This situation occurs most often in cases where an incentive program relies on small incentives being given to a large number of people. One example of this is the study context in chapters 3 and 4, where a coordinating entity is effectively precluded from offering monetary rewards because of tax implications of disbursing cash, and namely that the amount of information that would need to be collected from participants in exchange for a cash incentive of less than ten dollars a year make cash incentive effectively impossible.

A final issue with leaning on cash incentives is in the myopic view that this entails. The adoption of information technology has changed a number of facets of life – communication, transportation, education – among countless others. Motivation is among these. The use of information technology allows for rewards to be distributed with effectively zero marginal cost, sidestepping issues of transaction costs, with access controls which we theorize will avoid problems of motivation crowding. We place special emphasis on the “badge” construct, within the stream of gamification.

Gamification, the addition of elements first used in video games in non-game information systems, presents a solution to this. Video games often rely on motivational structures to engage players during intermissions between interesting aspects of the game or to increase replayability and play time out of titles. One such structure is that of the badge, which is an entirely optional reward which acts as a signifier of achievement.

Unlike cash, these badges have no instrumental use or transferable value to the player, which may reduce their susceptibility to causing motivation crowding. From a cost perspective, beyond the fixed cost of implementing the system, marginal costs of awarding badges are negligible.

“Badges” in the general sense constitute a broad range of implementations. At their most basic level, a badge is simply an icon which gets displayed to a user. Beyond this, further additions support adding context to the badge: accompanying text which describes the badge, a name for the badge, and so on. Unearned badges, especially in gamification schemes, are typically displayed with text which notifies the agent of the success criteria. More complex implementations exist, such as in game systems which display a meta-score computed on earned badges.

While we acknowledge the vast universe of possibilities that badges may be drawn from, we focus ourselves on badges which are composed of an image, and which includes success criteria which are known to the user. This format is the most common among gamification systems badge implementations. Though, even in this form, a variety of badge options are available, some important ones being studied within this dissertation. Indeed, badges are recognized to represent a broad multidimensional concept rather than a single incentive instance. Goals can differ, the format of the badge can change, social aspects can be added in terms of different mechanisms for displaying or not displaying badges to others, to name just a few.

Furthermore, beyond the components making up a badge, the behaviors which they are attempting to motivate are varied and often complex. For instance, within this dissertation, we are attempting to motivate commuters to choose to ride a bicycle rather than drive. This is a complex behavior for which incentives can be equally complex. The goal of the research is not to detail the full mechanisms by which incentives can operate in these settings. Instead, the goal is to induce noticeable changes in behavior with a small alteration in the motivational environment in the form of badges. The goal is in in the form of investigating the role of badges as a decision nudge, as described by Thaler & Sunstein (2008). These decision nudges are small environmental changes that help individuals bring their short-term decisions in line with their long-term commitments.

The following studies are exploratory in nature and show interesting, if not conclusive, results. These studies lay the groundwork for future research within these domains. As such, each chapter concludes with future plans for the research stream.

The following three chapters seek to evaluate badges as replacements for cash with the goal of avoiding negative aspects of cash incentive. The first, chapter two, deals with the problem of motivation crowding in a lab setting. The following two chapters, three and four, seek to use badges as a replacement for low-value, high-transaction cost compensatory schemes.

Chapter 2 - Badges, Motivation Crowding, and Transferability

Motivation is typically categorized into two discrete groups: *intrinsic*, which comes from within the agent, and *extrinsic*, which is external to the agent. Each of these sources of motivation comes with a set of properties that are theorized to have significant impacts on the downstream effects of the motivation (Benabou & Tirole, 2003). Intrinsic motivation, by virtue of its origin within the agent, is often durable while extrinsic motivation is fleeting and likely to evaporate as soon as the motivator is withdrawn.

To a first approximation, this categorization has a great deal of face validity. However, the recent explosion in popularity of gamified information systems calls this classification into question. The gamification of information systems overlays elements that originated within video games onto information systems that are not intended for entertainment usage. Common elements implemented include points, badges, and leaderboards, or “PBL” (Liu, Santhanam, & Webster, 2017a). For the purposes of this study, we focus on badges, icons awarded by the information system upon the satisfaction of certain criteria. These badges are undeniably an extrinsic motivator: they come from outside of the agent with success criteria determined by processes beyond the agent’s control. At the same time, the mechanisms that badges operate on are closer in nature to intrinsic motivators, driving value by enhancing agents’ perceptions of themselves rather than by offering transferable value that an agent may exchange for goods that satisfy other motives.

This transferability is a crucial differentiator among extrinsic motivators, e.g., distinguishing badges and cash. While symbolic awards exist that have similar self-perception enhancement effects, information systems are unique in their ability to offer virtual awards with no transferability. A trophy – perhaps the closest tangible analogue to a badge – may derive most of its value to the recipient as an indicator of achievement, but the recipient can still transfer the physical object to another. This transferability confers a nonzero additional benefit to the award. Cash, on the other extreme end, derives its value primarily from its transferability, even if the cash is a component of an award recognizing behaviors that may otherwise have reinforced a person's self-perception.

Indeed, cash remains to be one of the most common award mechanisms and extrinsic motivators used in practice. However, cash as a motivator has some well-understood limitations. One of these is that it typically results in the *crowding out* of motivations that might be present otherwise. These crowding-out effects occur when an agent's own intrinsic motivation is sufficiently high that the addition of extrinsic motivators causes the total motivation to exceed what is required, or *overjustified* (Wrzesniewski et al., 2014). This overjustified motivation causes the agent's intrinsic motivation to be reduced, or specifically, crowded out. This reduction in intrinsic motivation often persists even after the removal of the extrinsic motivator. As such, cash incentive schemes, not carefully deployed, can be counterproductive to the goals of the principal.

Because cash as a motivator is both external in its source and highly transferable, it is not clear whether the crowding effects will generalize to the use of badges as motivators. Badges are also external in source but are not transferable generally. We conjecture that the unintended consequences in the form of crowding out that arise from extrinsic motivation are a result of the transferability of the reward, rather than the origin of the reward. Previous research has focused on the origin of the motivator (e.g., Frey & Jegen, 2001) rather than the transferable properties of the reward itself. The goal of this study is to leverage this non-transferability to determine whether extrinsic motivators, deployed in a non-transferable fashion in the form of virtual badges, can avoid motivation crowding.

Literature Review

In this section we review the extant literature on intrinsic and extrinsic motivation, including the phenomenon of motivation crowding. We discuss both forms of motivation, bringing out a confounding that has existed within their characterizations. Separating the two different aspects of extrinsic and intrinsic motivators is important in considering gamification elements as motivators, and badges in particular. We use the following nomenclature throughout to discuss the two aspects of this distinction: *internal* and *external* will describe the locus, or source, of the motivation with respect to the agent. Particularly in discussing badges as motivators, it is important to distinguish the locus of the motivation from a second factor that has sometimes been conflated with locus as a feature distinguishing extrinsic and intrinsic motivations, namely transferability. *Transferable* and *non-transferable* describe the ability of a motivator – or certain properties thereof – to be given to another person. This distinction separates cash, an external transferable motivator that has been shown to create crowding out of other, intrinsic motivations, and badges, an external non-transferable motivator.

Motivation: Sources and Transferability

Extrinsic Motivation

Extrinsic motivation occurs in the form of external rewards or regulation (which can be conceptualized as avoidance of negative rewards). Benabou & Tirole (2003), for example, list rewards that are doled out by an external entity as extrinsic. These extrinsic rewards align with economic concepts of incentive and motivation based on the promise of external rewards. The notion of external incentives to motivate behavior in an agent[1] is axiomatic in economics to such a point that the ability of external incentives to motivate is not debated so much as the amount of incentive that would be required. Neoclassical economics, based in the theories of rational behavior, hold that extrinsic reward is neutral to internal motivation. Adam Smith, in the *Theory*

[1] Per convention, I will refer to the agent with the generic *he* and the principal with the generic *she.*

of Moral Sentiments, holds that the "lower" – i.e., commercial – virtues dominate the "higher" virtues, which map approximately onto modern conceptions of intrinsic motivation (Herzog, 2011).

However, in focusing on cash as a prototypical extrinsic motivator, previous research on sources of motivation collapse the source and the mechanism into the single dichotomous construct of intrinsic vs. extrinsic motivation (e.g., Cameron & Pierce, 1994; Deci, Koestner, & Ryan, 1999; Ryan & Deci, 2000). A defining characteristic of an extrinsic motivator is its source, i.e., that the motivation comes from outside the agent. However, another key attribute of cash as an extrinsic reward is that cash is highly transferable in nature: Most or all of the value of the reward can be transferred from the person who originally obtained the reward to another person without loss in value. Thus, cash as an extrinsic incentive collapses two motivational features. First, the source is external to the agent; the cash is provided generally by an external principal to influence the agent's behavior. Second, the cash as incentive has an externality that is easily transferable to another in its entirety. This transferability is a critical part of the award from an economic viewpoint. However, these two features of an extrinsic incentive do not necessarily conjoin. We return to this idea after considering the counterpart of intrinsic motivation.

Intrinsic Motivation

Early conceptualizations of intrinsic motivations revolve around the satisfaction of basic biological urges, namely: satisfying hunger and thirst, seeking out sex, and reducing pain (Clark L. Hull, 1943). These "tissue needs" – so-called because of their bypassing of the nervous system – arose out of observations of animals, though later researchers noted the wide range of human activities that do not directly satisfy any tissue needs – consider Maslow's hierarchy of needs as a prominent example of this (Maslow, 1943). For example, humans (as well as some animals) manipulate and explore objects despite not satisfying any tissue needs through these activities (Deci & Ryan, 1985).

Contemporaneously, Deci (1975) described intrinsic motivation as arising from activities where there is no reward but the completion of the task itself, explicitly connecting intrinsic motivation to the internal *source* of the motivation. This connection to an internal source of the motivation, where no principal is providing external rewards at the principal's discretion, is consistent with a definition of intrinsic and extrinsic motivations in terms of their respective sources. However, just as with extrinsic motivation, this clear-cut definition of intrinsic motivation in terms of having no external source is confounded by the mechanisms by which intrinsic motivations might operate. Thus, intrinsic motivations are sometimes defined, not in terms of the source of the motivation, but in the mechanism or form of the motivation. The mechanisms involved in intrinsic motivation generally relate to the idea of self-determination, wherein agents' behavior is characterized as coming from within the self. Thus, self-determination is consistent with the current characterization of intrinsic motivation as being partly defined by the internal source of the motivation. Tying to mechanisms, Deci identified self-determination as a higher level of motivation where tasks enhance one or more of three primary components: autonomy, or a sense of not being controlled into performing a task; mastery, or the sense of performing a task well; and relatedness, or a sense that performance of a task engenders a sense of connection to others. These self-determinative aspects are held out as a means by which motivations can operate. In terms of our distinction between transferable and non-transferable incentives, the aspects of self-determination comprise mechanisms by which non-transferable incentives operate.

Furthermore, autonomy, as a component of self-determination, is tied closely by Deci & Ryan (1985) to intrinsic motivation as defined as having an internal source. As such, it is an important concept in researchers' understanding of intrinsic motivation and, specifically, of the crowding-out phenomenon whereby intrinsic incentives are subverted by the addition of external rewards and regulation. Deci & Ryan write:

> To be truly intrinsically motivated, a person must also feel free from pressures, such as rewards or contingencies. Thus, we suggest, intrinsic motivation will be operative when action is experienced as autonomous, and it is unlikely to function under conditions where controls or reinforcements are the experienced cause of action (p. 29).

In other words, there is a somewhat mixed claim here. First, there is a claimed connection of autonomy to the operation of motivations whose source is internal to the agent; and, second, there is a possibly claimed mechanism via an *experience* of autonomy.

Here, we disentangle this confounding that is often present within discussions of incentives by distinguishing (1) between incentives with an external or internal source; and (2) transferable or non-transferable incentives, further distinguishing three aspects of non-transferable incentives: autonomy, mastery, and relatedness. Though we make this distinction, it is worth noting that external sources and transferable rewards, and internal sources and nontransferable rewards, frequently go together. It is not by accident that the confounding of these two dimensions has occurred in characterizing extrinsic and intrinsic motivation. Cash is nearly always employed as an external reward, and the transferability of cash is a defining feature of its use. On the other side, feelings of mastery, that might provide incentive when no externally sourced award is provided, cannot be transferred.

While research in this area is sparse, previous examples in literature do acknowledge the differences between internal and external motives and the mechanisms on which they operate. For example, (Wrzesniewski et al., 2014) study motivating factors of West Point Cadets with differing motives, all of which are external: family pressures, professional advice, and instrumental motivation. The first two are external in nature but operate on motives that are typically categorized as intrinsic; the latter is classically extrinsic. Notably, this dichotomy also holds for transferability.

Beyond this, with information technology, counterexamples have expanded. Agents can utilize reward structures that are created by themselves and are transferable. “Beeminder”, for example, is set up by the user and deducts cash from the user’s bank account if he fails to meet internally specified goals. Conversely and importantly for the purposes of our study, badges are externally derived but operate on an agent’s sense of mastery, a non-transferable mechanism. So, with this theoretical framework and its distinctions in hand, we are in a position to consider the similarities and differences between cash and badges, as a precursor to hypothesizing the effects of badges on motivation crowding.

Transferability as an Unexplored Construct

As noted, cash as a motivator combines two distinct features: It is an external source of motivation, it is a transferable form of motivation. In terms of its transferability, cash occupies an extreme position along a continuum. In contrast, consider a trophy as an award. Like cash, clearly its source is external; it is provided by an external source based on external criteria, outside of the agent being motivated. However, whereas the physical object itself is transferable—it is trivial for a person to give the trophy to another person—the achievement which the trophy represents cannot be transferred. More generally, rewards that symbolize accomplishments or positive attributes will be resistant to full transfer since the symbolic components of the award are closely tied to the task for which they were awarded. These symbolic properties do not transfer when the physical trophy exchanges hands. In contrast, cash as an incentive is generally external and fully transferable with little or no symbolic value. As such, cash generally does not necessarily trigger any of the non-transferable mechanisms characteristic of self-determination theory and potentially seen with other motivators. The cash, with its transferable value, provides its own motivation.

Compared to cash, badges, as examples of IT-enabled gamification awards, are also external. One does not typically award a badge to oneself. Unlike cash, badges have symbolic value and are virtual awards. As such they have little or no transferability; badges exemplify an instance where the source of the motivation and the mechanism of the motivation are distinct from the usual relationship observed with extrinsic and intrinsic motivators as discussed in the literature and noted above.

First, consider the shared feature of having an external source. Even for a motivator with an external source, the motivation must be internalized as an influence upon the agent to perform a task. With outside-determined motivations, the agent experiences some combination of *introjection*, or an acquiescence to value imposed on the agent without accepting the value for himself, and *integration* where the value is accepted by the agent as a part of his value system (Deci, Eghrari, Patrick, & Leone, 1994). Cash is perceived as involving aspects of both processes. Due its transferability, the value of the money is transferred from the principal to the agent, implying an introjection on the part of the principal. The money in itself has value to the agent, as well, thereby being readily integrated. Gamification strategies, like badges, operate primarily through integration, trying to activate values of the agent to motivate the agent's behavior, while not transferring any value from the principal. The form of this activation is hypothesized to be quite different given the difference in transferability between these two forms of external motivation.

Cash, being fully transferable, generally has a value in and of itself that is accepted by most people. It is a value that is within individuals' existing value system. Badges, like other gamification elements, do not have inherent value since they do not share this feature of transferability. Therefore, the mechanism by which they operate is presumably to activate non-transferable, self-determinative aspects of motivation, and specifically the experience of mastery. Specifically, the badge necessarily must have a goal condition that, when achieved, enhances the agent's sense of mastery of the task at hand. At the same time, the lack of transferable value on the part of the badge sidesteps the possibility of the introjection of other values, as the lack of transferable value precludes any use of the motivator for non-symbolic purposes. This combination of features opens the possibility that badges support an agent's non-transferable motivations in a way that avoids the consequences arising from the instrumental qualities of a transferable motivator, e.g., motivation crowding.

Motivation Crowding

One consequence of using money as a motivator is a motivation crowding effect. Most often, this is characterized in terms of the motivation source, as an external motivation crowding out, i.e., decreasing, internal motivations. The effect makes some intuitive sense: if performing an action requires a minimum motivational level of i, already present within the agent, and e external rewards are added, the agent can reduce his own motivation by e while still maintaining sufficient motivation to complete the task (Frey, 1997). This leads to problems when the removal of the extrinsic reward does not necessarily cause the internal motivation to return to its previous level. In fact, numerous empirical studies indicate that in general, it will not (e.g., Charness & Gneezy, 2009; Gneezy & Rustichini, 2000). Thus, motivation crowding involves two related phenomena. First is that adding an external motivation creates a substitutionary effect; the motivation does not add to the internal motivation, rather the internal motivation is replaced when it is not needed. Second is that the internal motivation does not automatically rebound when the external motivation is removed.

In order for these phenomena to occur, the introduced motivation must be taken as having precedence to the agent over the motivation that it is supplanting. The introjection of the external motivation must be internalized, i.e., accompanied by a new motivation that supplants the existing motivation. Being an external motivator is not necessarily sufficient, introjection does not need to be accompanied by integration. The external motivator can be perceived as a form of coercion, an infringement on one's autonomy with no connection to the agent's own values. Thus, in the case of cash as an incentive, the crowding out phenomena must arise chiefly from the transferability of money as a motivator. This transferability gives the money value to the agent, integrating the motivation into the existing value system of the agent. This integration can create a situation in which the behavior is overjustified (Wrzesniewski et al., 2014), leading to the result of the substitutionary phenomenon of crowding out, i.e., trimming off unneeded motivation. Still to be explained is why transferable incentives should take precedence when they are introduced and drive out the existing motivations that are non-transferable.

One suggestion, as to why money should have a greater strength as an incentive, allowing it to take precedence and substitute for already existing non-transferable motives of mastery, arises from the theory of relational models (Fiske, 2004). According to the theory, relationships between people are governed by four primary modes which generally describe the internal models that people use to interpret, construct, evaluate, and seek social relationships. Two of these relational models are equality matching (EM), which is most common in relationships where interactions are expected to be repeated, though not so close to rise to the level of familial intimacy, and market pricing (MP), where no further interaction is assumed. Both forms of interaction among individuals describe a transactional view of relationships. MP relationships are typified by the finalized settling of accounts after each transaction, while EM relationships are characterized as approximating equality in the long run. In contrast are Communal sharing relationships, which involve a sharing of benefits based on needs and a concern for joint welfare, and not based on exchange with an expectation of reciprocation; and, Authority Ranking, where social interactions have a superordinate/subordinate structure. These forms of interaction among individuals describe non-transactional views of relationships.

These four relational models are conceptualized as membership categories (Haslam, 1994). Thus, a particular social interaction can have different levels of membership in multiple of the four models simultaneously; and, the bringing to mind of a relational model is a categorization task based on social cues. In this way, people use transactions to infer information about the relationship. For example, the offer on the part of one person to pick up the lunch tab in exchange for reciprocity on the next outing implies an EM relationship. As such, the manner and the nature of exchange offers are significant in forming the perception of the mode of relationship.

As noted, money is a transferable incentive. As such, it signals a transactional social view, as observed with MP and EM relationships. More specifically, transferable objects, by their nature, have discoverable value. The combination of this feature and the offer of a transferable reward itself signal an MP relationship, since the value transfer can be made without repeated interactions between agent and principal.

This shift in perception, whereby the introduction of transferable incentives tends to signal a MP mode of social interaction, creates a feedback loop that strengthens the transferrable incentive as a motivator compared to non-transactional incentives like a desire for mastery. Badges, being non-transferrable, are less likely to engage an MP viewpoint, thereby lessening the likely perception of the exchange as instrumental. As such, we expect that the motivation will be perceived as primarily supporting the internal, non-transactional motivation for the task, thus sidestepping the possibility of substitution and crowding.

Properties of Badges

Though non-transactional, badges do contain within them certain inherent properties which are inseparable from the badge itself. As noted earlier, offering a badge as an award requires a goal condition to be known to the participant. Having goals, even without rewards, has been shown to be motivating on its own: simply presenting participants with a goal can increase the likelihood that they will meet that goal (Locke & Latham, 2002). Furthermore, having a goal generally also presumes some mechanism by which the agent is able to monitor progress toward the goal. Previous research in motivation crowding has demonstrated that making the monitoring measurement salient to participants can cause crowding-out (Etkin, 2016). As such, while our primary interest is between using a badge and not using a badge, we also are interested in which of these aspects of badges are responsible for the incenting properties of badges and, if badges do lead to crowding, which is responsible for the crowding. We add two additional manipulations to separate out these effects. Since the awarding of a badge in the experimental task is time-based, we add the following conditions: (1) a timer-only condition, the timer being the mechanism for monitoring progress toward the goal; and (2) a condition with a goal and a timer, but no badge.

We posit that as a visible symbolic award for performance, the badge will enhance the mastery incentives for a task beyond any motivational benefits that a timer and goal can provide without the badge, or in other words, that the badge reward itself has motivational power. While the literature suggests that goals, even those imposed by outside parties, can be motivating, the reward (in this case, the visual indicator) is the proverbial carrot which provides the primary motivation. Anecdotally, this has support from within games as well: before badges, games would commonly include “par” times without specific rewards; only with the advent of badges, and in games with badges, did cultures develop around these features.

Hypothesis Development

The nature of a reward is determinative of its motivational effects. For a reward with economic value that provides a transferable incentive, the economic value will dominate any intrinsic, non-transferable value that the award might have provided. As a result, cash as an incentive leads to motivation crowding. The introduction of information technology allows restrictions on transferability within a system; and badges, by their nature, are generally non-transferable. At the same time, badges are undeniably external: they originate from within the information system, not the agent. Whereas traditional reward schemes are frequently internal and non-transferable or external and transferable, we leverage the unique aspects of IT to create rewards that are extrinsic and non-transferable. These mechanisms can be “bolted on” to commercial information systems, or such affordances already exist and are under-utilized.

Because of this non-transferability and thus minimization of perceived economic value, we conjecture that the agent will not perceive extrinsic motivators in the form of badges to be substitutionary in nature, but rather to enhance mastery as a feedback mechanism, which is an internal, non-transferrable motive. As such, we expect that the badge will enhance the agent’s sense of self-determination and thus his intrinsic motivation.

H1: *The offer of a badge as incentive will increase performance relative to: (a) not receiving a badge, (b) receiving only a timer, and (c) receiving a timer and a goal but no badge.*

At the same time, since the badge is hypothesized to increase the user’s sense of intrinsic motivation, the agent will not suffer from the negative consequences of motivation crowding as discussed in Frey (1997). Being non-transferrable, the badge is hypothesized to support the existing motivation, not supplant it. Without crowding out and substitution of motivation, agents' perceptions of their own intrinsic motivations are still present when the badge is removed, both within themselves and as perceived by others; thus:

H2: *The crowding out phenomenon will not occur with: (a) an offer of a badge as incentive, (b) receiving only a timer, nor (c) receiving a timer and a goal but no badge.*

In addition to testing Hypotheses H1 and H2 with game performance measures, keystroke information is collected during the study. Previous literature suggests that increased activity in computer inputs indicates higher levels of arousal (e.g., Lee, Tsui, & Hsiao, 2015). The game involved successive decisions of aiming a shot and then firing a bubble with each shot, as indicated in the Methods section below. These actions utilized three keyboard keys: the left and right arrow keys to orient the firing direction of the bubble, and the up-arrow key to fire the bubble. The keystroke data were recorded and transformed to generate two additional metrics. First is the number of keypresses on each level, measured on a per-minute basis. This metric is intended to capture the rapidity of keyboard inputs on the part of the user. The second is the number of lateral movements per shot (left + right / up). These keypress-based metrics proxy for the player's motivation to complete the game by measuring the quantity and quality of inputs. Since the game rewards both speed and accuracy, the relationship between the two measures is of particular interest relative to the study's hypotheses. Finally, we collect data on the amount of time taken to complete each level, as the primary goal of the game is to complete each level in a minimum of time.

Methods

Participants

The study was conducted on the Amazon Mechanical Turk platform over the course of six weeks with no participant filters in place. 285 participants completed the game task and survey and were compensated $0.75 each. Of the 285 participants, 19 were discarded due to anomalous gameplay data; specifically lack of lateral movement of the firing mechanism (described below) within one or more levels, indicating a failure to engage meaningfully with the task. Furthermore, we include attention check questions (see Appendix 5) and discard participants who did not indicate a notice of the manipulations, which led to 15 records being discarded. For the remaining 251, mean time to complete the survey was 724 seconds (12.07 minutes), for a normalized compensation of $3.728 per hour, which is roughly in line with Amazon Mechanical Turk norms (Buhrmester, Kwang, & Gosling, 2011).

Context

We utilize an open-source game, *Frozen Bubble*[2], which has been used in the past for experimental purposes (e.g. Liu, Li, & Santhanam, 2013). This game is attractive as a test platform for studying crowding out phenomena for a number of reasons. First, it is a game, and is intended to be enjoyable in its own right, and as such should involve greater levels of *ex ante* intrinsic motivation compared to other tasks common to the motivation literature. Because it is a game, there is a degree of mastery that can be attained, and the game itself lends itself to perceptions of improvement. Finally, the game's mechanics are reasonably simple to understand, even among people who are not avid video game players.

The game interface, in its basic form, appears as shown in Figure 2-1. Mechanically, the game is a clone of *Puzzle Bobble* (later *Bust a Move*), wherein a player fires bubbles of assorted colors at a field. The firing mechanism is at the top of the igloo at the bottom of the display with a colored bubble loaded for firing. The firing mechanism can be angled to different degrees to the left or right using the left and right arrow keyboard keys. Once the angle is set, the colored

[2] Original (Perl) repository is here: https://github.com/kthakore/frozen-bubble; Code used in experiment is available upon request.

bubble is fired in a straight line using the up-arrow keyboard key, bouncing from the sides if the bubble's path is unrestricted. The bubble's progress stops when it contacts one of the bubbles in the central area of the game display. If the point of contact is with one of at least two contiguous bubbles of the same color, the game applies rudimentary physics wherein all the contiguous bubbles of that color, and any other bubbles that are otherwise disconnected from the top of the display, fall and disappear.[3] If the bubble that is fired does not initially make contact with one of two or more same-colored bubbles, it remains in the central field at the point where it made contact with a bubble. Increasing difficulty, the top of the central display, from which the bubbles are hanging, moves down periodically during the level, decreasing the playing area (further screenshots can be seen in Appendix 2). The goal of each level is to clear out all the bubbles in the central part of the game's interface before the bubbles reach the bottom of the display. Doing so allows the player to proceed to the next level. If the bubbles reach the bottom of the central area before the level is cleared, then the player loses the level.

The game was heavily modified from its original form, emphasizing completing levels quickly in the new version, whereas in the original, the primary goal was to clear out all bubbles without time being a factor. Given the change in goal, the notable change made to the game rules was the addition of a timer. Furthermore, although failing a level would still be registered as a failure, the player moves to the next level whether successful or not. These changes move the perception of mastery from the completion of the level (as in the original game) to the speed of completion, with badges being awarded on each level for completing the level within a given time limit. These changes were made for time consideration, as the original game, played masterfully, could take several hours to complete.

[3] A video of this can be seen at https://z.umn.edu/FrozenPhysics

Figure 2-1: Basic (control) interface

Furthermore, the original game underwent significant modification by using a re-written version using JavaScript rather than Perl (which was the original language used) or Java (which was the basis for the game in Liu et al. (2013)). Effort was made to keep individual design choices intact; however, it is important to caution against over-interpretation of results across different game versions.

Procedure

We randomly assigned participants into one of four groups, with each group adding an additional feature to the design, as follows:

- **The "control" condition** provides no information to the participant.
- The **"timer"** condition places a timer on the left sidebar of the play field.
- The **"goal"** condition adds a specified goal time beneath the running time, calibrated to be attainable by approximately 50% of participants based on preliminary tests of the game design.

- The **"badge"** condition adds a badge for the successful completion of the level before the goal time, indicated by a gold star under the goal indicator, one for each of the three levels.

Instructions shown to each treatment group are included in Appendix 2.2.

Participants in all four conditions are directed to a study page on a University server where they provide their informed consent and are then presented with instructions on playing the game, including an animated image demonstrating the game physics. Text for these are included in Appendix 2.2. Participants are directed to use sound and a computer with a physical keyboard. Before continuing, participants must explicitly confirm each of these.

After consent and instructions, the participant plays the game in "practice" mode without any manipulations, comparable to the control condition of the study. Because the game physics make up a significant mechanic that is critical to success, acclimation in a low-stakes environment allows the participant to understand the game before manipulations are introduced. This procedure is consistent with other tasks which have a learning curve (cf. Baumeister 1984). Two levels were presented for practice.

After the practice period, participants are presented with information introducing them to the game that varies based on the manipulation (see Appendix 2.2). Four levels of manipulations were used, each additive to the previous: control, timer, goal, and badge (detailed below).

A primary dependent variable occurs after the three game stages. The participant is informed that data need to be uploaded and verified by our servers, which is expected to take a few minutes. They are given the option of playing a few more levels ("bonus") or waiting on a blank screen. The choice made by the player at the end of the game stage – to play the bonus stages or wait - is recorded. The degree to which participants are not interested in continuing to play the game constitutes a measure of crowding out.

After the bonus (waiting) period, participants are presented with a number of questionnaires for use as experimental checks and controls, as detailed in the next subsection (see Appendix 2.5 for the text of the items). Finally, a code is furnished to the participant to submit to Amazon Mechanical Turk to verify completion.

In addition to these procedural elements, other data were collected without direct user intervention. These include keystroke information, the amount of time spent on each page and to complete each level, as well as basic information sent from the browser which include the web browser name and version, operating system, and similar.

Controls

The text of the questions related to variables examined as controls and checks in the analyses is shown in Appendix 2.5. Participants are sorted into conditions randomly. We assume *ex ante* that the randomization process is sufficiently robust as to distribute unobserved factors across conditions, even considering sampling biases that are common with Turk (Paolacci & Chandler, 2014). Turk is admittedly a convenience sample, so we check that the constructs under study do not vary across demographics that are unique to Turk's peculiarities. Industry-backed market research finds more interest in video games among persons under 40 than above, consistent with the approximate age of the onset of "digital natives" (Prensky, 2001), and further that 41% of self-identified video gamers are female (Entertainment Software Association, 2017). We collect information on the age and sex of the participants to check for equal distribution among the conditions.

Given that the dependent variable involves playing a game, participants who play video games frequently may have higher *ex ante* levels of intrinsic motivation with games and experience greater levels of motivation crowding. As such, we ask participants to indicate how often they have played video games in the past month.

Also included in the post-experimental survey is a seven-item shortened Intrinsic Motivation Inventory (IMI; Deci & Ryan, 2003) used as a robustness check against the primary variable with the logic that over-justification produces lower average levels of intrinsic motivation (Frey & Jegen, 2001). The shortened version of the IMI minimizes time spent in the survey while maintaining construct validity. The shortened scale measures interest, enjoyment, and perceived competence, while discarding subscales orthogonal to the present research such as “Value/Usefulness” and “Relatedness”.

Three questions similar to the intrinsic motivation inventory but framed around users’ feelings about the game were also asked and separately validated. These questions were included to provide further opportunities for robustness checks. However, the scale produced a Cronbach’s alpha of .66, which we did not consider to be sufficient to indicate internal validity. As such, analysis of the feeling questions was abandoned, and they are not discussed further.

Results

Randomization Checks

Participants were assigned treatment conditions randomly upon acceptance of informed consent. Distribution of treatment conditions are shown below:

Control	**Timer**	**Goal + Timer**	**Badge + Timer + Goal**
59	60	67	65

Mean age of respondents was 33.21 (s.d.: 9.63) years and the mean experience was 3.8 days (s.d.: 2.28) of gameplay per week. Of the 266 participants, 97 were female (36%), which is not significantly different from the Electronic Software Association's 41% (2-tailed proportions test with continuity correction, p = .36). As these factors are ordinal, we test using the Kruskal-Wallis (KW) rank-sum test, wherein none of these factors significantly differed across treatment groups: game experience ($KW\ \chi^2(3) = .801, p = .849$), age ($\chi^2(3) = 4.376, \mathrm{p} = 0.224$), or sex ($\chi^2(3) = 0.125, \mathrm{p} = 0.946$))

Outcome: Choice to Play

We use a number of outcome variables to measure motivation. We begin with a look at hypothesis H2, the prediction that badges will not lead to crowding out. The primary dependent variable is the player's decision to continue playing when given the opportunity. Across the four treatment groups, 28.2% of participants chose to play rather than wait. Figure 2 shows the by-group data, the differences across the groups were not statistically significant ($\chi^2(3) = .761, p = 0.517$) consistent with hypothesis H2.

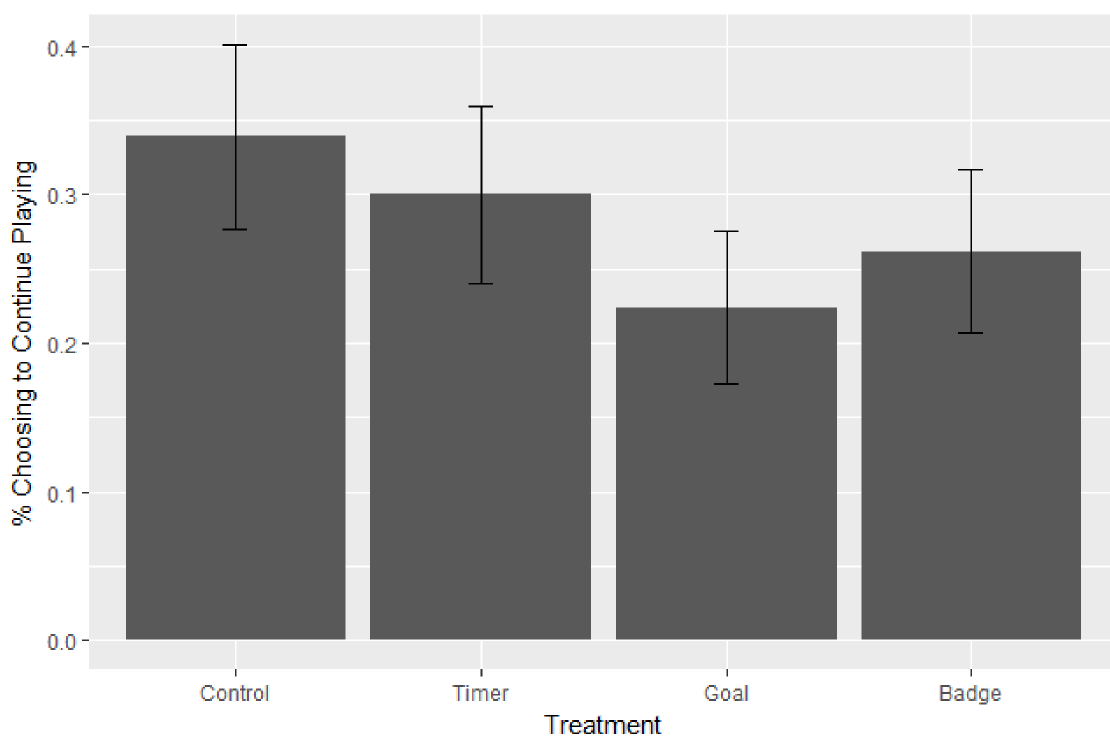


Figure 2-2: Percent Continuing to Play (error bars represent standard error)

As an additional measure of a possible crowding out of intrinsic motivation, intrinsic motivation was measured after the waiting period, using a short form of the Intrinsic Motivation Inventory (IMI) from Deci and Ryan (1994), a well-validated measure of intrinsic motivation. The IMI was adapted slightly to reference the game task. Within the study, the 7-item scale produced a Cronbach's alpha score of .94, indicating good adherence within the study group. (See Appendix 2.5 for the questions). This measurement also did not produce results significantly different across treatment groups using the Kruskal-Wallis (KW) rank-sum test (K-W $\chi^2(3) = 4.348, p = 0.226$).

As such, we cannot infer that the introduction of timers, goals, or a badge caused a decrease in motivation to continue playing. The failure to find a result for both measures is as expected by hypothesis H2. However, we cannot accept the null hypothesis, so the absence of a demotivating effect on its own does not demonstrate avoidance of the crowding out effect. A critical aspect of this is the crowding out of intrinsic motivation by extrinsic motivation. As such, we can bolster the result by attending to the motivation within the main game task of the experiment following the manipulation and prior to the waiting period. Since it is impractical to take intrinsic motivation measurements during the study, we use alternative methods for inferring intrinsic motivation during the game.

Motivation within the Game

All participants were presented with a practice task, which was specifically noted to be low-stakes and did not include any manipulations irrespective of the treatment group. The instructions presented to participants can be seen in Appendix 2.1. This served two purposes: first was to allow participants to learn the game and its mechanics. Second was to measure the performance of players when not exposed to manipulations, thereby providing a baseline for individual differences. We measure motivation within the game using three key behavioral components. We can capitalize on the within-subjects design to compare performance during the practice rounds, prior to any manipulation, to performance during the game rounds, when the manipulation is in place. Thus, for each measure we analyze:

Δ = Game performance – Practice performance

as the dependent variable of interest, testing for differences across the treatment groups.

The first behavioral measure is the actions per minute, which is a normalized measure of the number of keystrokes that were input, or the quantity of inputs. All else being equal, we expect more motivated individuals to perform a greater number of actions per minute. However, no treatment groups show a significant increase in the delta of actions per minute in the game mode over practice mode ($\chi^2(3) = 0.093, p = 0.964$).

The second behavioral measure is actions per shot, collected as a proxy for the quality of inputs and defined as the number of adjustments (left and right) on average that a player makes between each shot. This provides a tension with the game time; and, since the game time is the primary objective, we expect to see fewer actions per shot among people who are experiencing higher motivation to perform well as fine-tuning shots runs counter to performing quickly. In this case, analysis shows a difference between groups in the delta for actions per shot ($\chi^2(3) = 3.764, p = 0.011$). As illustrated by Figure 2-3, the primary effect that badges have upon actions per shot as a motivational measure is in the adding the monitoring of performance as an aspect that a badge provides (treatment 2 vs. treatment 1).

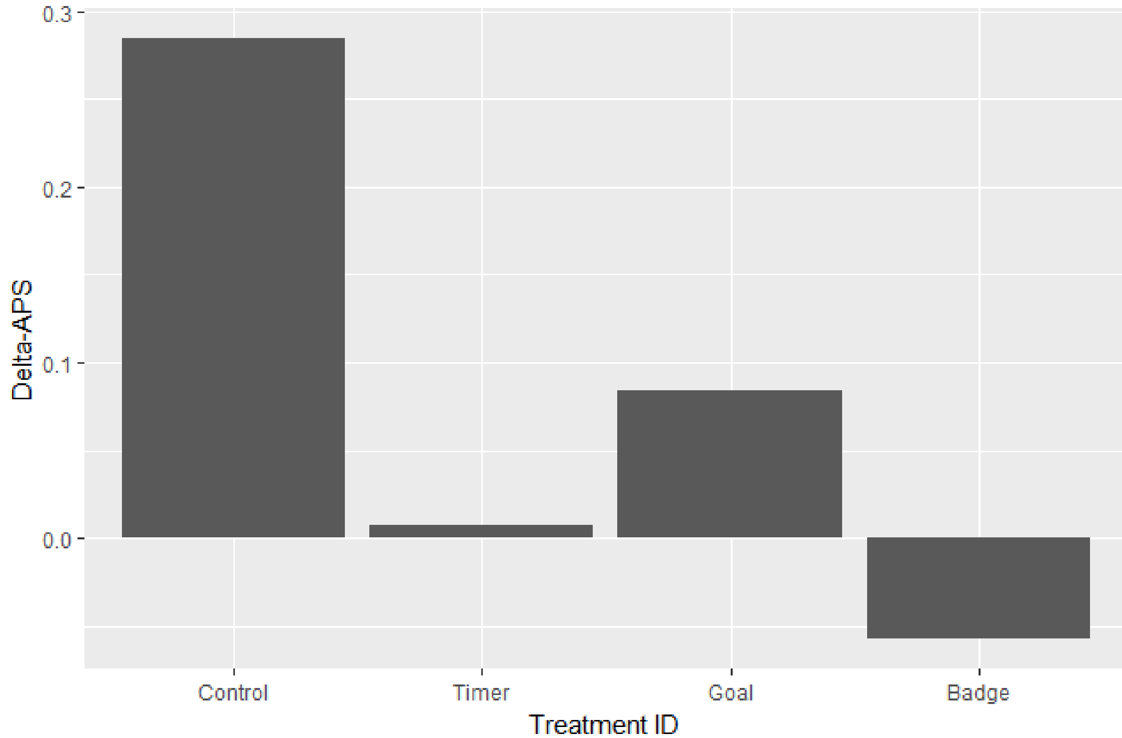


Figure 2-3: Change in Actions per Shot by Treatment ID

The third behavioral measure of motivation within the game is the average time per level: a greater delta for this variable indicates an increased motivation to complete the levels faster. Again, since time is the relevant variable for gaining a badge, decreased time per level is an indication of greater motivation. However, the delta for time per level did not significantly differ across conditions ($\chi^2(3) = 0.729, p = 0.536$).

Accordingly, the support for hypothesis H1 is mixed at best. While the actions per shot indicator point to a treatment effect, the total number of keystrokes and the time taken to complete the level do not. Further, it bears mentioning that while the actions per shot metric is not without value, actions per minute and the time taken to complete the level are arguably more indicative of motivation in this task as the speed of inputs and time taken to complete the level are both directly tied to the game goal, whereas the quality of shot is a second-order effect. As such, it remains unclear if the badge incentive increases motivation to perform the task (hypothesis H1). We develop the implications of these findings, in particular the indicated directions for additional study, in the following, concluding sections.

Discussion

The task chosen – playing a video game – was deliberately selected because games are designed to be deliberately intrinsically motivating, and thus a good platform to test for the possibility of badges crowding out intrinsic motivation. The results of the experiment provide no evidence to indicate that the addition of badges crowd out intrinsic motivation out as has been observed with cash incentives. Across all conditions, differences in the choice to continue playing as well as the self-reported intrinsic motivation are statistically insignificant. In this, the results are consistent with the lack of an effect for badges, as stated in hypothesis H2.

However, in addition to hypothesis H2 being a null hypothesis and unable to be directly supported, the other analyses also raise difficulties with drawing a clear conclusion about the crowding that badges may or may not induce. In order for motivation to be crowded out, we would need to see an increase in motivation with the introduction of badges, as described by hypothesis H1. However, the evidence for this motivational effect of badges, in the present design, is mixed at best. While there are some effects that present themselves by treatment group – namely actions per shot, which indicate a motivation to finish quickly by foregoing fine-tuning of shot placement – other indicators do not support an increase in motivation of introducing badges. The problem here is not in badges per se not being able to provide motivation. Other research, including as described in the other chapters of this dissertation, suggest otherwise. Instead, the issue is connected to the particular task and experimental setup being used. This leads directly to the consideration of future work, specifically to a strategy of altering the experimental setup and task to explore the issue further.

Before doing this in the next section, it is also noteworthy that the effect of badges upon actions per shot was observed even when only the monitoring aspect of badges was introduced. The theoretical identification that badges entail several aspects, and that their cumulative effects should be considered, receives support. The aspects of a badge as a reward are a system that monitors progress toward a goal, the presence of the goal, and the badge award for achieving that goal. In the present study, the monitoring component was sufficient for the effect of badges upon actions per shot as a motivational indicator. This suggests that this theoretical breakdown of the badge as motivator should be considered in future work that investigates badges as awards, and other related gamification strategies, more generally.

Finally, given the equivocal nature of the results, the work currently has limited implications for practice. At the very least, there is a possibility that by using badges, an external motivator that removes the transferability aspect that cash incentives entail, there is a possibility that the negative side effect of crowding that occurs with financial incentives may not generalize to external incentives like badges that are not as transferable. If future research can bear this out, the finding would offer an important incentive option in a variety of settings. This invites further study of the crowding phenomenon in different contexts, a topic on which we now offer a few suggestions.

Future Work

In future work, we intend to follow-up on the gaps in this empirical work, specifically exploring the question of whether or not badges do increase motivation in a measurable way. While the research question – whether the avoidance of motivation crowding can be achieved by the use of nontransferable rewards – remains a valid one, the task chosen may not have been as intrinsically motivating as we had expected beforehand. 28% of participants chose to continue playing when the alternative option was to stare at a blank screen: while no similar study has been done with which to compare this, this number suggests that the task may not be as intrinsically motivating as we had previously hoped. Beyond this, a number of internal motivations are conflated which require further experiments that seek to separate specific motivational aspects. First, it is unclear if the badge incentive is failing to motivate users. One reason that this is unclear is the lack of a definite motivator, specifically cash. We plan to offer cash incentive for good performance in lieu of badges to determine the comparative efficacy of the badge incentive. Second, we plan to modify the task with the intention of comparing how the intrinsically motivating qualities of the task (as expressed through choice to continue playing) interacts with the ability of badges to motivate players as well as the effect of crowding, if any.

Finally, further research into the specific components of badges is warranted. Our analyses are unclear as to whether the measurement, goal, or reward are the primary locus of the motivation. Since the measurement and goal aspects of badges are difficult to extricate from the badge itself, locating the motivator, or motivators, within the badges remains an important question. As with above, modifying the task to one which is more motivating on its own may go toward addressing this gap.

Chapter 3 - Proximal and Distal Badge Goals

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, Santhanam, & Webster, 2017b). Badges have attracted interest from a number of scholars, investigating the effect of badges in contexts from purchase intention (Shang & Lin, 2013), to education (Sitra, Katsigiannakis, Karagiannidis, & Mavropoulou, 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 "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 the badges' goal conditions affect individual motivation 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.[4]

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 in a manner similar to symbolic awards (Kosfeld & 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

[4] 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.

fewer risks to pedestrians, and promotes increased health for the riders. However, the nature of the student population makes monetary incentive effectively impossible. We explore the use of badges to motivate bicycle ridership among students at the University of Minnesota.

To accomplish this, we investigate the effects of presenting 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”. “Challenge” goals are directed at those who do ride sporadically but who, with sufficient incentive, would become more consistent riders, referred to herein as “infrequent riders”. Each of these 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 low-power 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 & 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. Rather, the benefit of the symbolic reward is the reward itself and nothing more. This does not necessarily preclude rewards associated with the symbolic award that accrue as second-order effects of the award, for example increased status among a peer group. But, the value of badges is primarily symbolic, with no 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, Scott, & McKee, 2002), and performance in rote tasks such as data entry (Kosfeld & Neckermann, 2011). Previous work has also demonstrated positive effects of symbolic awards when offered *ex ante* as a token of commitment (Baca-Motes, Brown, Gneezy, Keenan, & Nelson, 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 definite criteria exist for 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 choice architecture and motivational psychology literature, to which we now turn our attention.

Goals

For the purposes of this study, we draw on the 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 (in contrast to “do-your-best”), challenging goals enhance a person’s motivation (Locke & Latham, 2002). Goals also act as reference points against which performance can be judged (Heath, Larrick, & Wu, 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 & 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. We conceive of goals as a one-dimensional line starting at the initial state and ending at the goal state, with the current state lying along the line according to the person’s perceived progress. 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 (C. L. 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, Urminsky, & Zheng, 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-andanswer 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 close to the end of a goal path exhibited greater levels of posting behavior relative to users who were closer to the midpoint. 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.

Finally, when considering participants who do not exhibit any riding behavior, the concept of decision inertia applies. Decision inertia is the tendency to repeat previous choices, even when these choices are suboptimal (Alós-Ferrer, Hügelschäfer, & Li, 2016). This results from dual-process cognition, wherein the automatic system uses previous decisions as a default (Strack & Deutsch, 2004), as well as a result of the preference for consistency. However, the majority of study in these areas have been focused on probability tasks: Alós-Ferrer et al. (2016) focus on the role of Bayesian updating on ball draws, for example. We contribute to this stream by considering behavioral interventions with the goal of supporting wellness behavior.

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. When no badge is present, there is no added incentive to break decision inertia and 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 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. The goal-gradient hypothesis would predict that the proximity to the goal provides a motivating influence. We hypothesize that for a significant number of non-riders the motivation of the proximal goal will be sufficient as an incentive to ride. 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.

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 (Gibson & Radner, 1937) and decoy effects in judgment (Scarpi & 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. We therefore posit two contrasting hypotheses for the joint effect:

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. Thus, we largely expect this badge to be ignored by the infrequent rider. And, 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. Consequently, 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 exclude this group from our hypotheses.

Methods

Context and Description of Technology

The University of Minnesota, in conjunction with Dero, a bike rack manufacturer, developed technology for monitoring bicycle commuting, called the ZAP 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 students, who opt-in to 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.[5] While there is no direct financial incentive, the University has deemed it to be in the interest of the community to increase

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

bicycle ridership among students to decrease automobile congestion on campus and the surrounding area. Furthermore, because of tax implications, the University cannot issue direct cash payments to students to incentivize behavior. As a workaround, students who met a minimum number of riding days (twelve, or approximately 3 days per week) would be entered into a 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, de Moor, Boomsma, & de Geus, 2007).

The ZAP 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.

The second component are “readers” (Figure 3-1) 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.

Figure 3-1: ZAP Reader at Downtown Exit Route

Field Experiment

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 (no encouragement) design. The one-ride incentive was presented in the form of a gold badge that was shown or indicated as unearned by the image of a closed lock, and the three-ride incentive was presented as a diamond badge or by a closed lock if unearned (see Appendix 2 for illustrations). The no/no 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. (See Appendix 3.1 for more detailed information.)

Badges were presented in two ways. The "no" conditions in the main 2x2 received a weekly email on Sunday presenting them with a badge (if earned), goal text, and instructions on how to earn a badge. All participants could also access the same information through a dashboard on the Dero website, which all participants had cause to visit at least one time to register their RFID chip. 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. (See Appendix 3.1 and Appendix 3.2 for comparison.) Participants were presented with 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.

Participants

The participants were students at the University of Minnesota who had opted into the ZAP program and had installed a ZAP tag on their bicycle. This was further narrowed to participants who had recorded at least one "ZAP" 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 were excluded from analysis. All users of the ZAP program who did not opt-out were included in the study for a total of 2,616 participants. Of these, 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 3-1: Distribution of Participants across Conditions

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

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 Minnesota weather patterns (where the study took place) roughly coincide with Thanksgiving and Spring Breaks, students often take their bicycles home at Thanksgiving and retrieve them after Spring Break. Ridership patterns appear to lend credence to this (see Figure 3-2). The eight-week window was chosen as it coincides with the remaining instructional time between Spring Break and the end of the Spring term.

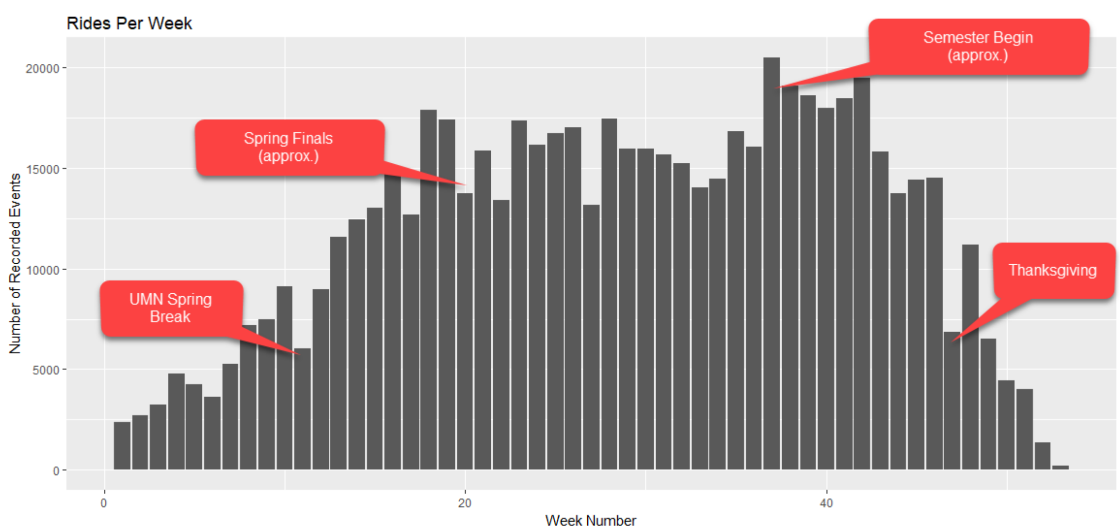


Figure 3-2: Ridership by Week of Year (2015-2017)

Since we divide hypothesized 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”. “Non-riders” were further subdivided into two groups, “zero” 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. These two low-riding groups are distinguished based on the underlying theory of hypothesis H1 based on the effects of decision inertia. It is expected that decision inertia likely will be even more applicable to the zero riders than for the rare riders, so they are analyzed separately. 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 3-2.

Table 3-2: Distribution of Rider Archetypes

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

To describe riding behavior within the group, please refer to the histogram faceted by rider type. For readability, zeros were removed (11 frequent riders, 58 infrequent, 154 rare, 1128 zero):

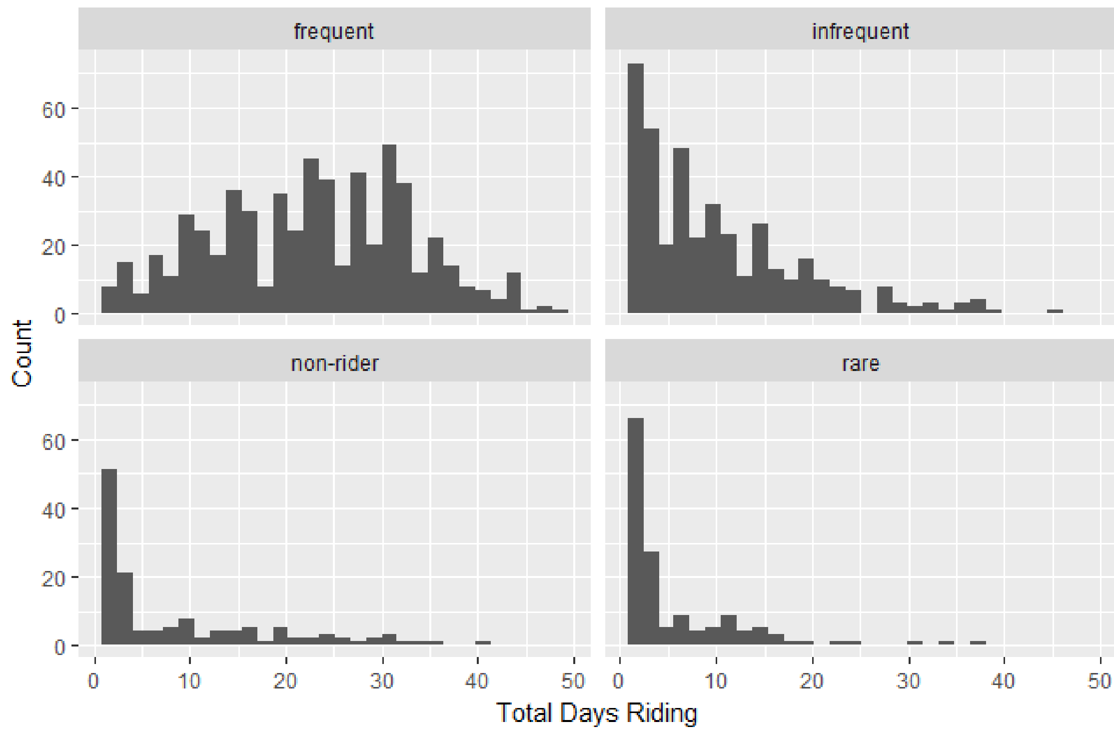


Figure 3-3: Histogram of Riding Days, Faceted by Archetype

Panels were created with (anonymized) user IDs, week IDs, the total number of zaps, and the number of unique days a zap occurred. We focus on days a zap occurred as placement of ZAP 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 zaps would imply more riding behavior. A person riding one mile across campus may record three zaps, 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 ZAPs rather has the potential to change riding behavior to maximize the number of scans rather than change behavior. 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 as a factor in our model:

$$days\ riding_{archetype} = \beta_{archetype,0} + \beta_1 participation_{archetype} + \beta_2 challenge_{archetype} + \beta_3(participation_{archetype} * challenge_{archetype}) + \epsilon_{archetype}$$

The input data for the model includes all behavior over the seven-week period arranged in a cross-section. The dependent variable, days riding, is the sum of all riding days over the seven-week study period. A riding day is defined as a day in which one or more ZAP event was recorded. 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. Model fit was evaluated with the Akaike Information Criterion (AIC).

Results

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 3-3. Weeks in Table 3-3 utilize Week 1 as a baseline and as such coefficients should be interpreted as change relative to behavior on Week 1. As indicated above, the model is fit separately for each archetype. As indicated in Table 3-3, none of the badge manipulations had an effect for this group, as expected.

Riders were binned according to their ridership behavior in the four weeks preceding the study period. Averages per group are shown in Table 3-2.

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 ($p = 0.089$). 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 ($p = 0.026$). 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 H3b.

For the infrequent rider archetype, the results were not fully consistent with the hypotheses. 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 ($p = 0.673$).

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

DV: Num. Rides over 7-week period	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
AIC	8758	3978	11179	18387

† p < .1,* p < .05. ** p < .01. ***p < .001.

Although not the subject of any hypothesis, the analysis for the frequent rider archetype group is included for comparison.

Discussion

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 and significant results, with zero-riders having a significant decrease in ridership ($p = 0.003$) and rare riders having a marginally significant increase ($p = 0.089$), and similar effects for the challenge badge: 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 supports an interesting pattern in terms of the offering of a participation badge to help break the decision inertia of being a non-rider and promote the onset of riding behavior. For the zero riders, not only did this not occur, but the indication was that offering a participation badge was a disincentive to riding. For the truly zero non-riders, decision inertia is indicated to be a strong force that is very difficult to break.

Badges alone may not be able to provide sufficient incentive to do so. Indeed, offering a participation badge when decision inertia is strongest may disincentivize the rider. With the rare riders, who would still be subject to some level of decision inertia but had already broken its effects at least to some extent, with a demonstrated proclivity to ride, these badges did appear to have efficacy, though only marginally so. Overall, 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 begin riding. Why, among non-riders, the participation it should serve as a disincentive is more difficult to explain. Perhaps not even being able to consistently ride once decreases motivation to ride even more.

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 that was a form of positive feedback, leading to an increase in ridership. 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 populations requires careful consideration of the target population, and then consideration of the idiosyncrasies of those targets with respect to the goal of the treatment.

Fortunately, this possibility is another capability that technology can bring to the situation. Being able to target different incentives to different individuals based on their past performance is a form of personalization for which technology is well-designed.

We also note that badges cannot be assumed to be monotonically beneficial, and that their deployment may come with adverse effects. The negative effects observed for zero riders represent an important consideration that badges may in fact have demotivational effects. Another case of this will also be seen in the next chapter of this dissertation.

The support for the expected lack of an effect on frequent riders also carries with it an implication for practitioners: that the addition of badges will have little effect on those who are already performing beyond the level which is being targeted. A concern for badge design is the possibility that adverse consequences will present themselves among non-targeted populations. This seems to not be the case among this group, who are already performing at a sufficiently high level and do not change this with the introduction of a badge.

Implications for the populations which are targeted present different challenges. Zero-riders did not respond to the incentive which indicates that the introduction of a badge was not a sufficient motivator. 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 while badges may not be without any effect whatsoever, their efficacy is likely to be reliant on the amount of behavior change which is required to earn them.

This presents implications for future research as well. First, it is clear 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 behavior change which badges can induce is still an open question and one worth exploring.

Chapter 4 - Socially-Framed Badge Design

The proliferation of video games from niche products primarily targeted to the socially awkward in the 1980s to common entertainment options that one's grandmother might regularly engage has been lucrative for the video game industry. At the same time, the evolution of these entertainment products to gain mass-market appeal also involved adaptation of new structures and mechanisms that motivate players to pick up or continue to play a game. However, despite the development of these structures within the context of games, their application is not limited to them. Thus was born *gamification*, or the use of game elements within a non-game information system.

Gamification implementations vary, but some elements find more frequent use than others. Points assign a numeric value to an action or behavior, communicating the designer's perception of valence and importance of an action. Leaderboards show a user's rank in terms of points (or some other countable metric) and allow for social comparison as well as competitive aspiration. Badges offer a goal and a token award for satisfying a pre-specified win condition. While there are other uses of gamification, points, badges, and leaderboards are sufficiently common as to earn an acronym, "PBL" (Liu et al., 2017b).

For the purposes of this study, we focus our attention on badges. First widely deployed with the Xbox 360 platform in 2005, competing game platforms subsequently adopted badge structures.[6] Badges offer goals that may not be critical to the completion of the game's main objectives, and in some cases may require playstyles that are remarkably different from the core mechanics of the game that they are associated with: for example, finding all the hidden objects in a level of an action game as a supplementary or secondary task to the central goal of winning the game. In

[6] We use "badges" generically; game systems often use unique names as a marketing strategy.

some implementations, badges contribute to a meta-game where a “Gamerscore” or level is computed across games based on the difficulty of earning the badges. Other times, some badges are unknown to the user until he earns them through the course of gameplay.

This is to say that while badges have a number of essential features in common with other implementations, the designer of a game (or information system) has to make a number of decisions on how to implement the badge. The importance of design holds at least as strongly for gamification applications, perhaps even more so since real-world consequences are typically involved. This motivates the broad research questions of interest in the current study: first, in a gamification application, do design features have an impact on the motivation of the user? And second, what arrangement of design features will be most effective in increasing motivation to perform the target behavior?

Outside of the context of video games, badges can be understood as a subset of symbolic rewards, or rewards which have no instrumental value to the awardee beyond the award itself (Kosfeld & Neckermann, 2011), or further, as positive reinforcement for behavior. This is a break from the study of incentive structures that are more commonly studied in economics, which presume that the value of awards is in their instrumental use (Benabou & Tirole, 2003), with money being the prototypical economic incentive. However, we understand that symbolic awards do indeed motivate people, even without the presence of incentives that can be used instrumentally by the recipient, both offline (Baca-Motes et al., 2012) and online (Hamari, 2017). Furthermore, symbolic rewards have a significant advantage to issuers of such rewards in their cost, which are frequently lower than cash incentives.

Another aspect of badges, shared with other gamification elements, is that they are online tokens intended to promote offline behavior. In this way, gamification extends the initial development of game awards that were online tokens to promote online activities. We use a novel field study context of a program within the University of Minnesota which rewards students, staff, and faculty for bicycle riding activity, measured in terms of days riding. Badges constitute a virtual, online reward to promote the offline activity of riding. Bicycle riding in the University context is an important behavior to encourage for three primary reasons: first, by encouraging bicycle commuting, the amount of stress that's put on the University's infrastructure is reduced in terms of traffic congestion, transit ridership, and parking needs. Second, bicycles do not emit pollution as cars do, and reduced automotive use increases air quality. Finally, bicycles do not constitute hazards to the pedestrian population as automobiles to. Bicycle riding has benefits that are limited to the rider as well as a result of engaging in wellness activity, such as increasing attention, reported well-being, and physical health.

Within this context, the goal is to explore the use of badges as a gamification strategy for motivating offline behavior with an online, symbolic award. The general question is how to optimally design these symbolic incentives in such a way that maximizes their efficacy in promoting offline behaviors.

Specifically, we focus on three design manipulations leveraging three forms of socially grounded incentives. The motivations behind these manipulations and their expected effects are detailed within the next section.

Literature Review

In this section, we begin with a review of badges and symbolic awards more generally. Then, we address the literature and hypotheses related to the specific design features explored in the study.

Badges and Badge Design

Symbolic awards are awards that have no instrumental value to the recipient beyond the symbolic value of the award itself (Kosfeld & Neckermann, 2011). Badges are a type of symbolic award which have no instrumental value to the recipient, but rather hold their value in the symbolism of their award. Previous literature has studied the efficacy (Hamari, Koivisto, & Sarsa, 2014) and optimal goal nature (Ch. 3) of badges. For example across varying settings, Baca-Motes et al. (2012) demonstrated the efficacy of symbolic awards in workplace contexts and (Hamari, 2017) similarly found benefits of using badges within video games. Since symbolic awards do not have any tangible value, their efficacy is presumed to arrive from their ability to activate or enhance an individual's internal motivation. Self-determination theory (Ryan & Deci, 2000) provides a lens into the means by which intrinsic motivation is affected by symbolic awards like badges. Badges in general symbolize the user's mastery relative to the goal to which the badge is attached, mastery being an internal motivation recognized within self-determination theory.

This efficacy of badges as motivational tools has been demonstrated when used as motivational tools in Learning Management Systems in terms of performing goal actions as defined by the instructor (Sitra et al., 2017). Badge design, however, varies considerably. Few elements exist that are intrinsic to badges: typically, an image, accompanying text, and achievement condition. Other aspects of badge design are not standardized across badge implementations. In cases where other elements, such as the ability to see other peoples' badges, have been studied, the effects on motivation to earn the badge are inconclusive (Hamari, 2013). An ability to see others' badges raises another major internal driver of motivation within self-determination theory, namely relatedness, which is a macro-theory of human behavior that describes the antecedents of intrinsic motivation (Ryan & Deci, 2000). Previous research has demonstrated links between compassionate and prosocial behavior and the relatedness construct (Hadden, Smith, & Knee, 2014) as well as finding relatedness as a key underlying motive for engaging in competitive behavior (Deci, Betley, Kahle, Abrams, & Porac, 1981). Applying this to badges as motivational tools, the general idea is that if badge design can be tied to socially-based features of the badges, the badge can increase motivation by activating relatedness motives in addition to the typical mastery motives to which badges naturally connect. Research addressing badge design features that can activate social motives of relatedness constitute a gap in the literature. Our study focuses on the use of socially-based design features that impact the incentive value of badges.

We direct our study toward three specific, socially based design features of badges: (1) warm glow – framing the badge award as a recognition of a prosocial act vs. for personal benefit, (2) social signaling – the ability to share the badge to Facebook vs. not being able to share the information to an online social community; and (3) a goal condition for which attainment is with respect to a fixed (non-competitive) goal vs. to a competitive goal. Each of these social components are features that have been considered as potential gamification features for promoting exercise activity and that have theoretical bases for being potentially successful as interventions. More specifically, each manipulation is intended to engender a sense of relatedness, the feeling of being related to others. Consequently, we expect a person's motivation to perform activities to increase when the activities are presented in a manner which highlights the person's relation to other people.

Framing: Achievement vs. Prosocial-oriented

Prior research on goal attainment suggests that specific goals are superior to non-specific, or "do your best" goals (Locke & Latham, 2002) that lack any external referent; however, more specific goals can be constructed in different ways. A key distinction, particularly in considering the potential of providing badges with social incentives, is between goals that are have an altruistic component, providing a warm glow, and those that are strictly achievement oriented. A “warm glow” is an increase in affect which is the result of “impure” altruism, in distinction to “pure” altruism, from which the actor derives no benefit whatsoever (Andreoni, 1989). This warm glow has value to the actor, and has been empirically found to increase altruistic behavior (Lilley & Slonim, 2014). This is largely because of the benefit of the warm glow to the agent. If tied to altruistic behavior, the badge acts as a symbolic reminder of the altruistic behavior performed by the agent, which the badge commemorates. From this, the badge provides the warm glow as a byproduct of the socially motivated altruistic act. Beyond the specific benefits to the recipient in the form of warm glow, performing altruistic acts leads to changes in neural activation (Harbaugh, Mayr, & Burghart, 2007), notably when the altruistic behavior is a choice rather than imposed on the giver. This increased affect is believed to arise primarily from an increase in the feelings of relatedness that are promoted by engaging in altruistic behavior. Consistent with this account, altruism generally increases self-reported feelings of relatedness. Thus, if the activity being rewarded has, or can be connected to, altruistic motives, then the enhancement of relatedness that is engendered is hypothesized to increase motivation for the activity.

In contrast, achievement-oriented framing recognizes that the motives are mastery-oriented with no component of promoting some social benefit. More so than altruistic awards, they can provide an additional incentive toward mastery. Since there is no natural conflict between achievement and altruism, the prosocial benefits of the warm glow are expected to add to the natural signal of mastery provided by a badge that is strictly achievement-oriented. However, the mastery orientation of recognizing achievement can also provide an enhanced incentive. This factor allows us to test the competing effects of enhancing relatedness vs. enhancing mastery. Since the badge itself in both cases has mastery implications, our a priori conjecture is that enhancing the new feature of relatedness will have a greater impact than further enhancing mastery motivations. Thus, we conjecture that framing the badge as an award for prosocial behavior will demonstrate greater efficacy as a result:

H1: Altruistic, warm glow badges will increase the target behavior over achievement-oriented badges.

Social Signaling

Humans are social animals, with strong needs to feel related to others and have social interactions (Kenrick, Griskevicius, Neuberg, & Schaller, 2010). Furthermore, people have a desire for their achievements and good deeds to be recognized and send signals to others for this purpose (Baumeister, 1998). Thus, adding affordances that allow users to share their badge awards with others on a social network is expected to increase the agent's motivation to perform the target behavior. The addition of a social network option foregrounds the relatedness aspects of the bicycling program, further boosting motivation to ride. In this case, the contrast condition of not being able to signal the badge does not provide a potential competing motivation. Thus, we frame the hypothesis as:

H2: The ability to share badges to social networks will increase ridership with respect to users not able to share badges to social networks.

Since both prosocial ties and social signaling are hypothesized to operate by activating motivations toward greater positive feelings of relatedness, the two factors are expected to interact in their effects on behavior. When combined with social signaling, the altruistic behavior can also provide social value in terms of signaling (Bénabou & Tirole, 2006). Users who share an achievement-oriented badge understand themselves to be broadcasting a personal achievement, while riders who share a pro-social badge understand themselves to be signaling their altruistic actions.

People show a particular desire to signal their altruistic actions to others (Ariely, Bracha, & Meier, 2009), even to the point of bragging (Berman, Levine, Barasch, & Small, 2015). As such, while we expect a main effect of social signaling as described by Hypothesis H2, we expect that there will be an additional positive effect under conditions where the rider is able to signal an altruistic award, rather than an achievement demonstrating mastery alone:

H3: The ability to share badges will have a greater impact on ridership if the badge conveys a warm glow and not just achievement, i.e. a positive interaction effect is expected between the effects of a warm glow and of signaling upon behavior.

Goal Condition: Fixed vs. Competitive

A fixed goal has a clear target that does not depend on the actions of others, e.g., engaging in three activities per week. In contrast, competitive goals make achievement dependent on the performance of others as well, e.g., being in the top 25% for the week. The fixed condition has the advantage to the rider of allowing them to know precisely how much they must do to earn a badge. While this is preferable from the standpoint of planning behavior to meet a goal, this may not be optimal for increasing ridership based on the tendency of people to over-perform in competitive situations. In contrast, the competitive badge has the disadvantage of not having a clear target that the rider can plan toward; but, although it connects to relatedness motives, it has the disadvantage of possibly lowering feelings of relatedness rather than enhancing them.

Prior research suggests that competition does tend to drive behavior in excess of what would otherwise be displayed by an agent. For example, Feng, Fay, & Sivakumar (2016), in studying online auctions find that increased competition leads to an increased propensity to overbid, or to pay more for an item than they would have been able to purchase elsewhere in a non-auction context for less. This is attributed by Heyman, Orhun, & Ariely, (2004) to “auction fever”, driven by what they term a quasi-endowment effect, or an aversion to loss of something that they conceptualize having at some point in the future. While the context is different, second-price online auctions are a competition between two or more people who seek to maximize their economic surplus by withholding information unless necessary. Furthermore, the possibility of a quasi-endowment effect remains a real possibility with a badge tied to socially competitive goals. In terms of self-determination theory, competition tends to activate a joint combination of internal mastery and relatedness motives. The incentive is relatedness-based in being driven by comparison with other riders, and with a sense of mastery in the comparison. Although it is possible that the comparison with others does not enhance relatedness, our initial hypothesis, based on the joint activation of motives is that framing the badge award as a competitive feature will increase ridership, and hence:

H4: Badge award conditions framed as competitive with other users will increase ridership over a specific, fixed condition.

This hypothesized tendency for competitive goals to activate favorable motives has been observed especially among people who are most likely to self-conceive as being in that top 25% (Liu et al., 2013). In contrast, fixed goals do not share this comparative social aspect. Furthermore, other work has found that high performers’ intrinsic interest in a task was enhanced by competitive elements, with a corresponding decrease among low performers (Epstein & Harackiewicz, 1992).

Accordingly, we expect that riders who self-conceptualize as being top riders will be more affected by this manipulation of adding a competitive goal. While the data collected do not allow us to measure the self-conceptualization of individual riders, we proxy for this with the amount of riding from the previous week and specifically whether the badge was earned. The argument is that those who earned a badge recently based on a competitive goal will take this as a signal that they are in the top bracket. As such, we expect the interaction between the previous weeks' badge award and the competition framing to produce a positive result.

H5: Badge awards framed as competitive will be more effective on riders who received the badge in the previous period.

In summary, we explore the effects of relatedness motives with badge feature manipulations that are hypothesized to operate in slightly different ways. All three operate in an environment where the receipt of a badge is presumed to have an inherent mastery component, indicating the achievement of some goal. The warm glow vs. achievement factor examines the effect of promoting a relatedness motivation as a contrast with enhancing the mastery motivation. The sharing feature offers an added relatedness incentive in contrast to no motivation enhancement. We also posit a hypothesis concerning the expected joint impacts of these two manipulations. The third manipulation offers a competitive framing that touches on both mastery and relatedness goals vs. a fixed goals framing that solely touches on mastery. It is also recognized that this effect may be nuanced , with the relatedness aspect being seen either positively or negatively depending on the self-perception of the rider as being in the top tier of riders or not.

Methods

We conducted a field experiment with participants randomly assigned into one of eight conditions forming a 2 (warm glow [yes/no]) x 2 (fixed/competitive) x 2 (social signaling [yes/no]) between-subjects design.

Participants

At the start of the study period, 4,235 riders had registered with the University of Minnesota ZAP program. We excluded riders who had not registered at least one ZAP during the academic year under the assumption that they are no longer associated with the University. Furthermore, we exclude from analysis any users who opted out of receiving emails during the experimental period. Finally, we limit our analysis to riders who had registered their ZAP account on or before the beginning of the study period. This resulted in a panel of 1,975 users observed over seven weeks for a total of 13,825 unique user-week pair observations. Users were randomly assigned to conditions being distributed as shown in Table 4-1.

Social: No	Competitive	Fixed
Warm Glow: Yes	230	262
Warm Glow: No	247	269

Social: Yes	Competitive	Fixed
Warm Glow: Yes	243	246
Warm Glow: No	231	247

Table 4-1: Distribution of Participants

Context and Description of Technology

The University of Minnesota, in conjunction with Dero, a bike rack manufacturer, developed technology for monitoring bicycle commuting, called the ZAP 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, as well as the riders, who reap the benefits not only of the rebate incentive, but also of the numerous physical and mental benefits that are associated with increased physical activity (McAuley et al., 2000).

After developing this technology, the program expanded to include students, who do not participate in the University's insurance plan. 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 decrease automobile congestion on campus and the surrounding area. Because of tax implications, the University cannot issue direct cash payments to students to incentivize behavior. As a workaround, students who met a minimum number of riding days (twelve, or approximately 3 days per week) would be entered into a lottery for a $10 gift card, donated by local businesses. Even among students who do not ride frequently enough to enter the lottery, students can use the technological infrastructure to measure their riding behavior.

The ZAP 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 or by request to students, staff, and faculty, though participation in the program for all categories is voluntary. Each RFID chip has a unique serial number that is associated with a profile within the information system.

The second component are readers (See Figure 4-1) that are placed in high-traffic areas, typically entry and exit points. 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.

Figure 4-1: ZAP Reader at Downtown Exit Route

Procedure

Data collection took place over a seven-week period, March 19, 2016 to May 6, 2016. This period spanned from the end of the University's spring break to the beginning of finals week. This time period was chosen for a number of reasons. First, due to the nature of the weather in Minnesota, a large portion of the school year is inaccessible to casual bicyclists without specialized gear and, in some cases, tires. Furthermore, a number of students bring their bicycles off-campus (e.g., to parents' houses) at Thanksgiving and retrieve them at Spring Break. A visual representation of bicycle trips (across all years) supports this (Figure 4-2).

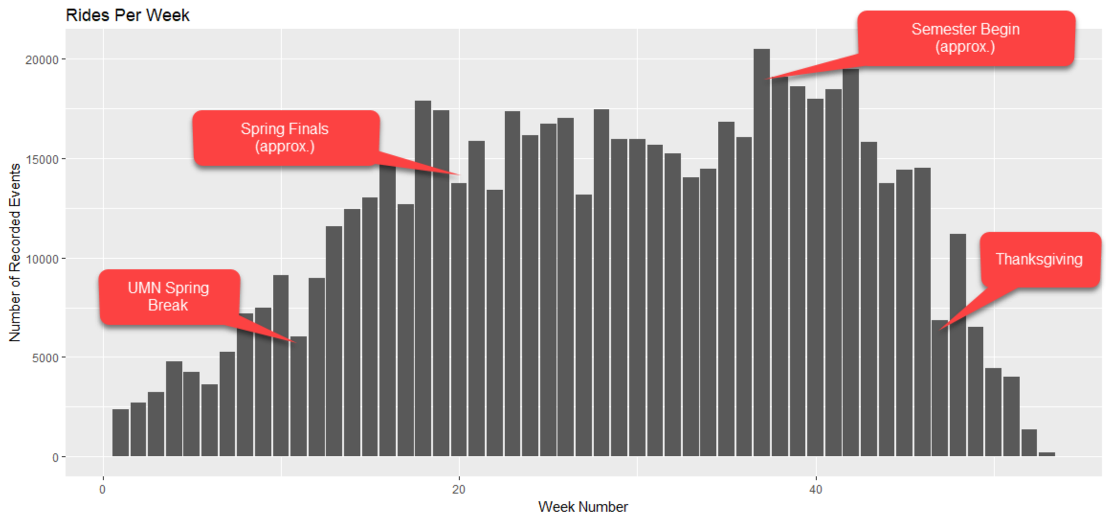


Figure 4-2: Ridership by Week of Year (2015-2017)

Each week, participants who had registered at least one ZAP in the preceding week were sent an email with a summary of their riding behavior as well as a presentation of their badge depending on condition and if awarded. Users who did not ride were not sent an email for that week. All users, including those who were excluded from the study or opted-out of email, had the option to log into the existing Dero ZAP dashboard where their badge for the preceding week would be presented, along with their riding behavior from the past week. See Appendix 4 for example emails and dashboards.

For all the eight different treatment conditions, all users received an email and dashboard display with the following common elements: the name of the user, their number of days riding, and an image and caption of the badge, if awarded for that week. The caption was titled “Star Rider” for users in the achievement-frame condition, and “Eco Healer” for users in the prosocial-frame condition. According to the framing condition, a different image was displayed, see Appendix 4.3. The badge was presented in color if they had earned the badge, and in black and white if they hadn’t. A line of text would indicate that they had (or had not) earned a badge, and included instructions on how to earn the badge, which would be presented as “Ride 3+ times to earn a badge" for the competition condition, and "Be in the top 25% of riders" for the fixed condition.

Users in the social condition were presented with a link to share the badge to Facebook, users in the non-social condition did not see further text. To prevent leakage, the link to share the badge would present the user with an apology and explained that the feature was under construction and logged the user's attempt. Seventeen users clicked this link.

The instruction text in the competitive condition notes that the badge will be awarded for being in the top 25% of riders. Users in the fixed condition, on the other hand, are told that a badge will be awarded for three rides. In fact, all users, regardless of condition, earned a badge for riding three days in a week. This approximates the previously used incentive scheme which required twelve days riding in a calendar month. We note that of riders in a given week, the 75^{th} percentile is typically three rides.

To summarize, the manipulations used for the 3 factors in the 2x2x2 between-subjects design were as follows:

- **Warm Glow/Achievement**: Users in the warm glow condition are presented a badge and text which emphasizes prosocial frame (Eco Healer badge), users I the achievement condition were presented with a badge and text which emphasized personal achievement (Star Rider badge).
- **Social Sharing (Link/No link)**: Users with the manipulation were presented with a link to share the badge and text to Facebook, users outside of the manipulation did not receive this.

- **Competitive/Fixed**: Users in the competitive condition were informed that being in the top 25% of riders would earn a badge; users in the fixed target condition were told that three rides per week would earn a badge.

Results

Across all 10,767 user-week pairs, 3058 badges were earned, or 28.4%, roughly equal to the 25% noted in the competitive condition.

We employ a logistic regression model using the log-odds of receiving a badge given the manipulations. Since manipulations are randomly assigned, we assume *ex ante* that they are exogenous. To avoid bias due to the incidental parameter problem, we use a pooled model rather than fixed effects. Furthermore, we utilize dummy variables for each condition: warm glow (1) vs. achievement (0), termed "warm glow" in the analysis, competitive (1) vs. fixed (0), termed "competitive" in the analysis, and social link active (1) vs. inactive (0), termed "social". We add an interaction term between warm glow and social to check hypothesis H3; and, we include an additional interaction term using the previous weeks' riding behavior $(badge_{t-1})$ to check hypothesis H5. In accordance with Hilbe (2011), we evaluate model fit with the Akaike Information Criterion (AIC). In mathematical terms, we use the following model:

$$\log - odds\ of\ badge_{it} = \beta_0 + \beta_1(warm\ glow) + \beta_2(social) + \beta_3(warm\ glow * social) + \beta_4(competitive) + \beta_5(competitive * badge_{it-1}) + \epsilon_{it}$$

All findings are summarized in Table 4-2. With respect to H1, a marginally significant difference ($p = 0.057$) was found in riding behavior between badges framed in a prosocial manner versus badges framed as a personal achievement. However, in support of H2, the effect of the ability to share badges to social networks does produce a statistically significant effect ($p = 0.011$). H3, positing an interaction effect from these two manipulations, does not produce a significantly supported difference ($p = 0.184$).

With respect to the effect of a competitive goal as captured by H4 and H5, we first note the statistically significant interaction effect on riding frequency of the rider earning a badge in the previous week and the competitive condition ($p < 0.001$) in support of hypothesis H5 and illustrated by Figure 4-3. The slopes are inverted among riders who earned a badge in the previous week with respect to the badge condition, indicating an effect of the interaction between an attainment of the previous week's badge and being in the competitive condition. As such, we find evidence for H5's notion that those who conceive as high performers are more likely to be motivated by competition, and conversely those who do not will be less motivated.

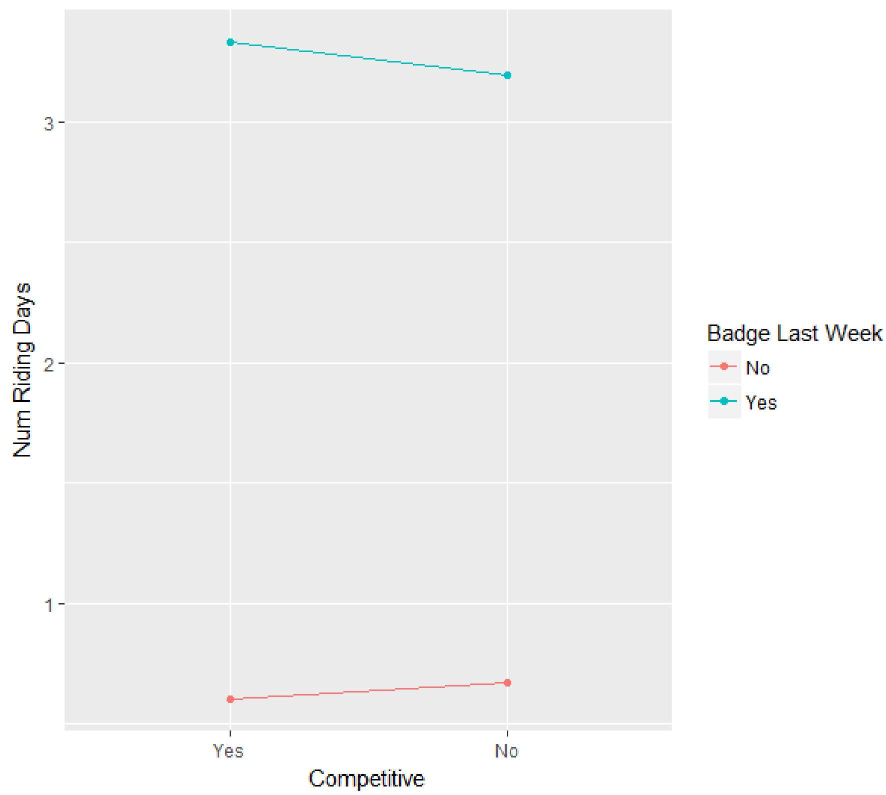


Figure 4-3: Interaction between Prev. Badge and Competitive

As clearly observed in Figure 4-3, despite that a statistically significant main effect ($p = .003$) shows for the competition manipulation (H4), this effect is clearly qualified by the interaction effect. So, we do not find support for hypothesis H4, though there is support that the previous behavior week's behavior is an indicator of the following week's ridership.

(DV: # of riding days/week)	Estimate	Std. Error	P (>|z|)	Sig.
(Intercept)	-2.433	0.064	< 0.001	***
Warm Glow	0.134	0.070	0.057	†
Social	0.179	0.070	0.011	**
Competitive	-0.192	0.065	0.003	**
Lag-Badge	2.990	0.071	< 0.001	***
Warm Glow * Social	-0.131	0.099	0.184	*n.s.*
Competitive * Lag-Badge	0.293	0.100	0.003	**

Table 4-2: Logistic Regression Model

(† p < .1, * p < .05, ** p < 0.01, *** p < 0.001), *n = 13823, AIC = 10863*

Discussion

The warm glow condition achieved only marginal significance in the model. Given the large *n* we hesitate to posit any practical significance and interpret this as indicating a lack of evidence to support the superiority of either the prosocial or the achievement framing. Accordingly, we fail to confirm hypothesis H1. In contrast, the social sharing ability leads to significant increases in days riding, confirming hypothesis H2. Both hypotheses were derived based on an expectation that adding a relatedness motivation to the mastery motivation that the badge represents, both constructs within self-determination theory, would increase performance. The difference in results connects to a difference in the nature of the manipulations of relatedness involved in each case.

The warm glow condition was contrasted with an achievement frame in the design. As noted earlier, the award of a badge for meeting a goal inherently ties to mastery motivations. The warm glow condition promotes added relatedness motivations, while the achievement manipulation enhances the mastery motivation in parallel. Thus, these two motivational enhancements may have operated in contrast with each other, neither dominating the other, leading to no support for H1. The social sharing manipulation was of a different character. The option to share the badge was contrasted with the lack of this option. The promotion of relatedness motivations by offering the sharing option did not have a parallel enhancement of the mastery motive in the no-sharing condition. In this way, this factor can be considered a purer manipulation of the impact of promoting relatedness, and hypothesis H2 is supported. Put another way, for the warm glow manipulation, mastery and relatedness were contrasting motivations in the two levels of this factor in the design, whereas for the social sharing manipulation, the added relatedness motivation was complementary to the mastery motivation provided by the badge.

Probably for similar reasons, the interaction between the warm glow condition and the social sharing condition, did not achieve statistical significance, failing to confirm hypothesis H3. This is presumed to be a product of the lack of effect on the warm glow condition, and further that advertising prosocial actions is not a more powerful motivator than advertising personal achievements.

Finally, hypotheses H4 and H5, relating to the effect of competitive framing on the users' rides, are considered together, beginning with hypothesis H5. The basis of H5 was the tendency for high-performing individuals to be more motivated by competition than lower performing individuals. The statistically significant interaction between the competitive framing and whether a badge was earned in the previous week lends support to hypothesis H5. Recall that in the competitive frame, it was expected that earning a badge in the previous week is a signal that the rider is in the top 25% of the population. And, previous research has suggested that such high-performing individuals are more motivated by competitive goals (e.g., Liu et al., 2013). The significant interaction in our data support this hypothesis H5. In addition, the inverse of this seems to also be true: low-performing individuals are demotivated by the prospect of competition.

Contributions

We contribute to the literature surrounding gamification in a number of ways. First, while the use of badges in a generic sense has been studied, the literature is less robust in optimal design of these features. We apply the relatedness construct from self-determination theory to offer insight in how badges may be optimized using well-studied behavioral mechanisms. Specifically, we apply framing devices around the nature of the reward in terms of different forms of relatedness-based incentives. While the prosocial frame was hypothesized based on the relatedness construct of self-determination theory, we note that mastery – the sense of performing well – is also a feature of self-determination, and in this context, it seems that neither dominates the other. When relatedness is pitched against mastery motives, for example in the warm glow conditions, neither dominates. Further research would need to disentangle the possibilities of whether this is due to the weakness of warm glow as a relatedness motive or the comparable, competing strengths of enhancing relatedness vs. enhancing mastery motives.

However, the positive impact of a sharing option, enhancing relatedness motives, does provide evidence for the general approach of using design features that incorporate prosocial elements as a means of enhancing the effects of badges in gamification settings. With a relatedness feature that does not have competing self-determination motives, we do, find evidence of efficacy.

The effect of introducing the relatedness aspect of competition also argues for a more nuanced view. Competition is an incentive that jointly impacts mastery and relatedness motives, as the introduction of competition both allows people to enhance their sense of mastery through competition, and competition with others triggers effects on one's sense of relatedness. For competition, the effect is not necessarily supportive, however, showing a statistically significant individual difference. As the people most likely to feel a sense of mastery when compared to others, we find that high-performing users will be motivated by the use of competitive framing. For them, the relatedness aspect enhanced the mastery component. However, with users who were not high-performing, the relatedness aspect had the opposite effect, leading to lessened performance.

In summary, the manipulations indicate an interplay between mastery and relatedness motivations when introducing badges as incentives designed to activate relatedness motives. Badges have an inherent value as a symbol of mastery: they are markers of achieving some goal. Relatedness can add to the benefits of mastery as with signaling, relatedness and enhancing mastery can have parallel effects (or non-effects) as observed in adding a warm glow vs. supporting achievement, or relatedness can prompt differing effects across individuals as observed when there is a competitive component.

Our research also has implications for practice. In our specific context, encouraging the specific behavior of bicycle commuting is of interest to transit managers, and the design of badges to encourage this behavior has broad application, especially given the task's ability to be perceived both as a marker of personal achievement as well as prosocial activity. This translates well to other applications that can be conceptualized both as a prosocial activity as well as an achievement or competition: amounts of recycling, collection for charity, and so on.

The use of badge design manipulations further informs the design of virtual symbolic rewards, especially within video games themselves, as the progenitors of badges as we use them, but also for other information systems designers. The inclusion of features that enhance the sense of relatedness provides a design template for designers of gamified systems, including learning management systems, fitness tracking, and the like.

Furthermore, the implication for competitive awards on the motivation of users who are not high performers should be considered when considering a feature of this nature. It is not strictly undesirable to cater to the high performers, though it is important to consider that those who are not high performers may experience demotivation in the face of a competitive reward.

Future Directions

This research also leaves a number of questions to be answered. The manipulations used represent a small subset of the possible applications of relatedness to badges, and future research in this area may expand this into other relatedness design choices. Furthermore, given the application of self-determination theory, a seminal paper in behavioral economics, to badges naturally leads to other behavioral mechanisms which may be inform the design of badges.

Several questions were not addressed by the experimental design. Foremost among these is the conflict between the warm glow and the achievement condition. It is unclear if the warm glow condition only weakly enhanced a sense of relatedness, or if both the relatedness motive in the warm glow condition and the mastery motive of the achievement frame were both equally strong motivators.

Finally, our experimental platform was limited to people who had indicated an interest in performing the behavior which was being encouraged. While this population made for a good group upon which to test badges, the boundaries remain unclear. Inducing wellness behavior in groups of people who are more reluctant to engage with the behavior, such as those in rehabilitation, may work to define the boundaries of how effective badges can be in promoting behavior.

Chapter 5 - Conclusions

The use of external motivators, and especially cash, are common for inducing agents to perform actions that they may not otherwise be willing or interested in performing on their own. While in many cases, this is not a problem – working for pay, for example – there are some cases where payment is used where a nudge may be more appropriate, or in cases where the use of cash payment may have negative effects. Among these are the problems of motivation crowding as well as transaction costs overshadowing the transaction-exclusive marginal cost of issuing the rewards.

We presented two cases where cash may not be the ideal motivator, with examples from theory and practice where cash fails, and offer an alternative with the use of badges. First is the case of motivation crowding, or overjustification, where cash, as a transferable reward, can have effects in the long-term that are contrary to the goals of the incentive scheme. While we do not find evidence to support the notion that badges act as effective motivators, we do not find evidence of their use causing an crowding out effect either.

Second is the case of high transaction costs relative to the marginal cost of the incentive. In our context, this high transaction cost was induced by regulation, but is likely to occur in cases where an entity has an interest in inducing small amounts of behavior change across a large population. In our case, this was a university transportation office seeking to encourage bicycling behavior among students. Badges in this case have the benefit of having a negligible marginal cost to deploy, but also fit squarely within a “libertarian paternalism” (Thaler, Sunstein, & Balz, 2010) ethos where the decision space is being altered in such a way which does not issue consequence to non-compliance. We find here that among subsets of our population, badges have an effect, and most often among those who require the least amount of marginal effort to attain the badge: most often, people who need to increase only nominally their performance of a behavior which they already perform (as with rare riders and badges), or among people whose more consistent performance of a behavior would earn them the reward (as with infrequent riders). This indicated that this effect is bounded, and that decision inertia is a strong force which may require more powerful motivators.

We conclude with design considerations for badges, using the relatedness construct within self-determination theory, complementing the mastery motive which badges inherently trade on. Relatedness is a core motivator of human decision making, and designing badges around this concept offers high leverage for information systems designers. Relatedness also operates on the aspect of self-determination theory which is not represented to some degree in badges inherently: the autonomy motive is baked in to the design inasmuch as badges by design have no penalty for non-acquisition, and the mastery motive is core to badges as well as other rewards, symbolic or otherwise. When these motives all align, we find that badges are at their most effective as motivating tools.

Notably, these are incentives that work more as “nudges” rather than strong motivators. This is a limitation of the use of badges, but not necessarily a strike against them. Providing inducement for small changes across large numbers of people may be infeasible with cash given transaction costs, but the aggregated benefit in a large organization can be considerable. Even for minimal changes to behavior, when these are spread across a large number of people, the effects can be profound.

We contribute to the research literature within this stream as well. First with the application of self-determination theory to the use of badges, and specifically within the context of motivation crowding and badge design. In both of these cases, we offer insight into the interactions between self-determination and decision making in the face of badge rewards. Furthermore, we test theories of goal gradient motivation and decision inertia to badges, finding that decision inertia is a powerful force, and that badges may not be effective motivators for goal-gradient motivation.

Finally, we offer some consideration for future research. First is the motivation crowding effect and badges, which we do not find conclusive evidence on given the questionable efficacy of badges on motivation in the assigned task. However, in later chapters we do find that badges have a motivating effect, which suggests that a study with different methodology may be able to more conclusively determine the link between motivation crowding and badges. Within the context of variable distance to badge attainment and decision inertia, the boundaries to which decision inertia overpowers badges is unclear, though the presence of an effect of decision inertia is suggested by our findings. Finally, the universe of badge design possibilities – even within the relatedness construct of self-determination – is vast and deserves future study. Information contained within the badge, information transmitted by the receipt (or non-receipt) of the badge, and abilities granted to the user to advertise their status with badges all operate on relatedness, among many others.

We re-emphasize the exploratory nature of these studies, though with an encouraging outlook. We have evidence which suggests the efficacy of badge design choices as well as a roadmap for future work. Badges are still a new concept in information systems. Their widespread use in video games is barely a decade old; the adoption of badges into non-game information systems has been in popular use for approximately half that time. Their continued study by researchers and employment by practitioners still has fruit yet to bear.

Chapter 6 - References

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

Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, *97*(6), 1447–1458. Retrieved from http://www.jstor.org/stable/1833247

Ariely, D., Bracha, A., & Meier, S. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, *99*(1), 544–555. https://doi.org/10.1257/aer.99.1.544

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

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*(3), 610–620. https://doi.org/10.1037/0022-3514.46.3.610

Baumeister, R. F. (1998). The self. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (2nd ed., Vol. 2, pp. 680–740). McGraw-Hill.

Benabou, R., & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, *70*(3), 489–520. https://doi.org/10.1111/1467-937X.00253

Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, *96*(5), 1652–1678. https://doi.org/10.1257/aer.96.5.1652

Berman, J. Z., Levine, E. E., Barasch, A., & Small, D. A. (2015). The Braggart's Dilemma: On the Social Rewards and Penalties of Advertising Prosocial Behavior. *Journal of Marketing Research*, *52*(1), 90–104. https://doi.org/10.1509/jmr.14.0002

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk. *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Cameron, J., & Pierce, W. D. (1994). Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of Educational Research*, *64*(3), 363–423. https://doi.org/10.3102/00346543064003363

Charness, G., & Gneezy, U. (2009). Incentives to Exercise. *Econometrica*, *77*(3), 909–931. https://doi.org/10.3982/ECTA7416

Deci, E. L., Betley, G., Kahle, J., Abrams, L., & Porac, J. (1981). When Trying to Win. *Personality and Social Psychology Bulletin*, *7*(1), 79–83. https://doi.org/10.1177/014616728171012

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating Internalization: The Self-Determination Theory Perspective. *Journal of Personality*, *62*(1), 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627-668; discussion 692-700. https://doi.org/10.1037/0033-2909.125.6.627

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4899-2271-7

Deci, E. L., & Ryan, R. M. (2003). Intrinsic motivation inventory. *Self-Determination Theory*.

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* (p. 763). New York, New York, USA: ACM Press. https://doi.org/10.1145/2470654.2470763

Ding, L., Kim, C., & Orey, M. (2017). Studies of student engagement in gamified online discussions. *Computers & Education*, *115*, 126–142. https://doi.org/https://doi.org/10.1016/j.compedu.2017.06.016

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *Academy of Management Review*, *14*(1), 57–74. https://doi.org/10.5465/amr.1989.4279003

Entertainment Software Association. (2017). *Essential Facts about the Computer and Video Game Industry*. Retrieved from http://www.theesa.com/wp-content/uploads/2017/09/EF2017_Design_FinalDigital.pdf

Epstein, J. A., & Harackiewicz, J. M. (1992). Winning is Not Enough: The Effects of Competition and Achievement Orientation on Intrinsic Interest. *Personality and Social Psychology Bulletin*, *18*(2), 128–138. https://doi.org/10.1177/0146167292182003

Etkin, J. (2016). The Hidden Cost of Personal Quantification. *Journal of Consumer Research*, *42*(6), 967–984. https://doi.org/10.1093/jcr/ucv095

Feng, C., Fay, S., & Sivakumar, K. (2016). Overbidding in electronic auctions: factors influencing the propensity to overbid and the magnitude of overbidding. *Journal of the Academy of Marketing Science*, *44*(2), 241–260. https://doi.org/10.1007/s11747-015-0450-9

Fiske, A. P. (2004). Relational Models Theory 2.0. In *Relational Models Theory: A contemporary overview* (pp. 3–25). https://doi.org/10.4324/9781410611413

Frey, B. S. (1997). *Not just for the money*. *Books*. Brookfield, VT: Edward Elgar Publishing. Retrieved from https://ideas.repec.org/b/elg/eebook/1183.html%5Cnhttp://ideas.repec.org/b/elg/eebook/1183.html

Frey, B. S., & Jegen, R. (2001). Motivation Crowding Theory. *Journal of Economic Surveys*, *15*(5), 589–611. https://doi.org/10.1111/1467-6419.00150

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. *Journal of Experimental Psychology*, *20*(5), 453.

Gneezy, U., & Rustichini, A. (2000). A Fine is a Price. *The Journal of Legal Studies*, *29*(1), 1–17. https://doi.org/10.1086/468061

Goette, L., & Stutzer, A. (2008). Blood donations and incentives: Evidence from a field experiment, (IZA Discussion Paper No: 3580), 1–34.

Hadden, B. W., Smith, C. V., & Knee, C. R. (2014). The way I make you feel: How relatedness and compassionate goals promote partner's relationship satisfaction. *The Journal of Positive Psychology*, *9*(2), 155–162. https://doi.org/10.1080/17439760.2013.858272

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. https://doi.org/10.1016/j.elerap.2013.01.004

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, *71*, 469–478. https://doi.org/https://doi.org/10.1016/j.chb.2015.03.036

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural Responses to Taxation and Voluntary Giving Reveal Motives for Charitable Donations. *Science*, *316*(5831), 1622 LP-1625. https://doi.org/10.1126/science.1245938

Haslam, N. (1994). Categories of social relationship. *Cognition*, *53*(1), 59–90. https://doi.org/10.1016/0010-0277(94)90077-9

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, *38*(1), 79–109. https://doi.org/10.1006/cogp.1998.0708

Herzog, L. (2011). Higher and lower virtues in commercial society: Adam Smith and motivation crowding out. *Politics, Philosophy and Economics*, *10*(4), 370–395. https://doi.org/10.1177/1470594X10386564

Heyman, J. E., Orhun, Y., & Ariely, D. (2004). Auction fever: The effect of opponents and quasi-endowment on product valuations. *Journal of Interactive Marketing*, *18*(4), 7–21. https://doi.org/10.1002/dir.20020

Hilbe, J. M. (2011). Modeling Count Data. In *International Encyclopedia of Statistical Science*. https://doi.org/10.1007/978-3-642-04898-2_369

Holgersson, H. E. T., Nordström, L., & Öner, Ö. (2014). Dummy variables vs. category-wise models. *Journal of Applied Statistics*, *41*(2), 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), 25–43. https://doi.org/10.1037/h0072640

Hull, C. L. (1943). Principles of Behavior: An Introduction to Behavior Theory. In *The Journal of Abnormal and Social Psychology* (Vol. 39, pp. 377–380). https://doi.org/10.1037/h0051597

Imai, K., & Goldstein, D. (2007). Nudging Turnout: Mere Measurement and Implementation Planning of Intentions to Vote. *Available at SSRN …*, 1–23. https://doi.org/10.2139/ssrn.977000

Intrinsic Motivation Inventory. (1994). Intrinsic Motivation Inventory (IMI). *The Intrinsic Motivation Inventory, Scale Description*, (Imi), 1–3. https://doi.org/www.selfdeterminationtheory.org

Kenrick, D. T., Griskevicius, V., Neuberg, S. L., & Schaller, M. (2010). Renovating the Pyramid of Needs: Contemporary Extensions Built Upon Ancient Foundations. *Perspectives on*

Psychological Science : A Journal of the Association for Psychological Science, *5*(3), 292–314. https://doi.org/10.1177/1745691610369469

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

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

Lee, P.-M., Tsui, W.-H., & Hsiao, T.-C. (2015). The Influence of Emotion on Keyboard Typing: An Experimental Study Using Auditory Stimuli. *PLOS ONE*, *10*(6), e0129056. https://doi.org/10.1371/journal.pone.0129056

Lilley, A., & Slonim, R. (2014). The price of warm glow. *Journal of Public Economics*, *114*, 58–74. https://doi.org/10.1016/j.jpubeco.2013.12.004

Liu, D., Li, X., & Santhanam, R. (2013). Digital Games and Beyond: What Happens when Players Compete? *MIS Quarterly*, *37*(1), 111–124. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02767783&AN=85634502&h=pylGYyQOVgRKZjcT4F9XB3xNrEz%2FIfTDdWPMcFXCrlteZG%2Bho0ko7bjNbpLRCYfNMpi1npOIZq0Nbf02JnPm7A%3D%3D&crl=c

Liu, D., Santhanam, R., & Webster, J. (2017a). Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. *MIS Quarterly*, *41*(4), 1011-1-A4. Retrieved from 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

Liu, D., Santhanam, R., & Webster, J. (2017b). Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. *MIS Quarterly*, *41*(4), 1011-1-A4. Retrieved from 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., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

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

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. https://doi.org/10.1037/h0054346

McAuley, E., Blissmer, B., Marquez, D. X., Jerome, G. J., Kramer, a F., & Katula, J. (2000). Social relations, physical activity, and well-being in older adults. *Preventive Medicine*, *31*(5), 608–617. https://doi.org/10.1006/pmed.2000.0740

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

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*, *23*(3), 184–188. https://doi.org/10.1177/0963721414531598

Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

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

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

Shang, S. S. C., & 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). Retrieved from http://aisel.aisnet.org/amcis2013/eBusinessIntelligence/RoundTablePresentations/11

Sitra, O., Katsigiannakis, V., Karagiannidis, C., & 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), 3037–3046. https://doi.org/10.1007/s10639-016-9550-5

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

Stubbe, J. H., de Moor, M. H. M., Boomsma, D. I., & de Geus, E. J. C. (2007). The association between exercise participation and well-being: A co-twin study. *Preventive Medicine*, *44*(2), 148–152. https://doi.org/10.1016/j.ypmed.2006.09.002

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. (C. R. Sunstein, Ed.). New Haven: New Haven : Yale University Press.

Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2010). Choice Architecture. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1583509

Wrzesniewski, A., Schwartz, B., Cong, X., Kane, M., Omar, A., & Kolditz, T. (2014). Multiple types of motives don’t multiply the motivation of West Point cadets. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(30), 10990–5. https://doi.org/10.1073/pnas.1405298111

Chapter 7 - Appendix

Appendix 2.1: Practice Game Instructions

Game Instructions

You will be playing a video game that involves clearing all bubbles on screen. If the bubble you fire hits a patch of same-colored bubbles, it will break them off. Otherwise, it will attach to existing bubbles. Bubbles will bounce off the sides of the stage.

You can adjust firing direction using **LEFT** or **RIGHT** keys on your key board, and **UP** key to fire bubbles. The compressor at the top will press down periodically. Once any bubble hits the bottom, the game is over. See the screenshot below.

You will next practice the game for two levels.

Play Practice Levels

Figure 7-1: Frozen Bubble Instructions

Appendix 2.2: Instructions

Instructions: Badge Manipulation

Instructions

You will now officially test the game with three new levels. Please click the "Next" button when you are ready.

During official testing, you will see a **timer**, which tracks how much time has elapsed in the current level. Above the timer is a **goal time**, which represents a "par" time. You can now earn a **badge** for each level (up to three badges), if you complete the level within the goal time

Next

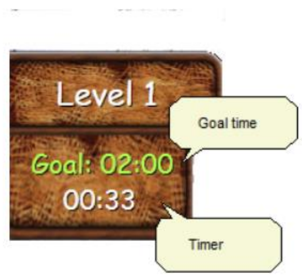
Figure 7-2: Badge Manipulation Instructions

Instructions: Goal Manipulation

Instructions

You will now officially test the game with three new levels. Please click the "Next" button when you are ready.

During official testing, you will see a **timer**, which tracks how much time has elapsed in the current level. Above the timer is a **goal time**, which represents a "par" time.



Next

Figure 7-3: Goal Manipulation Instructions

Instructions: Timer Manipulation

Instructions

You will now officially test the game with three new levels. Please click the "Next" button when you are ready.

During official testing, you will see a **timer**, which tracks how much time has elapsed in the current level.

Next

Figure 7-4: Timer Manipulation Instructions

Instructions: Control Condition

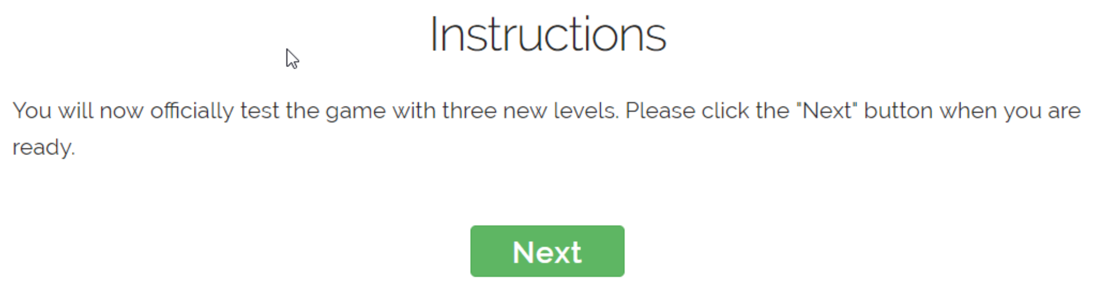


Figure 7-5: Control Condition Instructions

Appendix 2.3: Game Screenshots

During Play

Figure 7-6: Screenshot during Play

Success

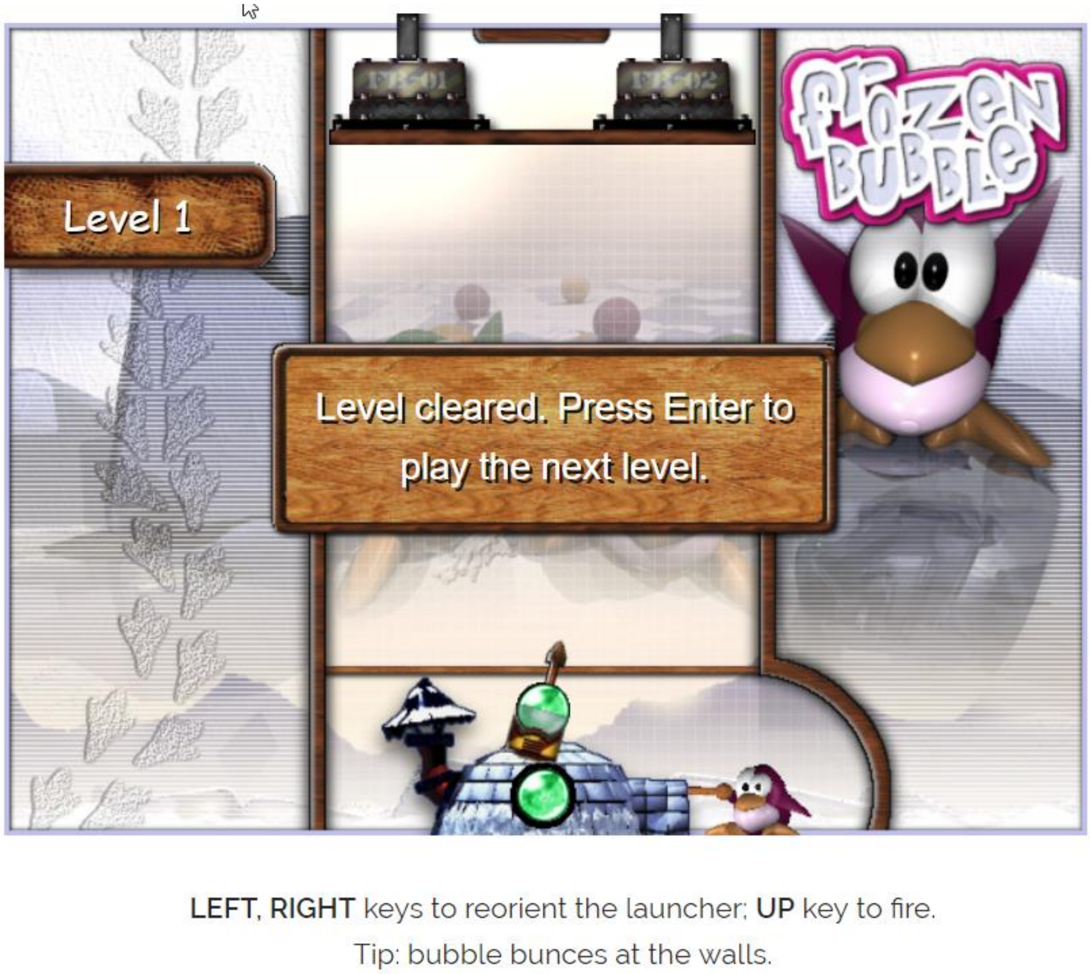


Figure 7-7: Success Screen

Failure

Figure 7-8: Failure Screen

Appendix 2.4: Choice Screen

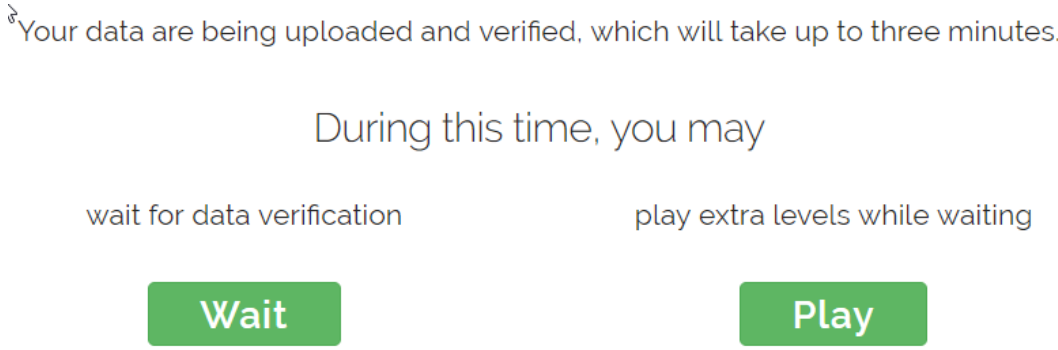


Figure 7-9: Choice Presented to Play or Wait

Appendix 2.5: Survey Questions

Intrinsic Motivation Inventory

1 (Strongly Disagree) – 7 (Strongly Agree) scale

- "I enjoyed playing this game very much ",
- "This game was fun to play."
- "I thought this was a boring game."
- "This game did not hold my attention at all. "
- "I would describe this game as very interesting. "
- "I thought this game was quite enjoyable. "
- "While I was playing this game, I was thinking about how much I enjoyed it. "

"Feeling" Questions

1 (Strongly Disagree) – 7 (Strongly Agree) scale

- "I was motivated to finish the levels quickly."
- "I was interested in playing even after the experimental period."
- "I enjoyed playing the game."

Manipulation Checks

(Yes/No)

- "Were you aware of the timer on each level?"
- "Were you aware of the goal time on each level?"
- "Were you aware of the badge incentive on each level?"

Appendix 3.1: Dashboard Images

"Gold" (Participation) Goal

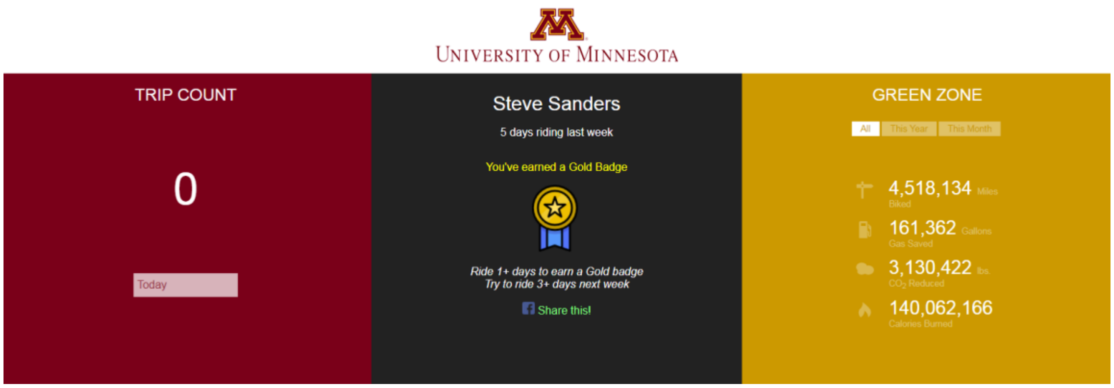


Figure 7-10: Dashboard with Participation Badge Earned

"Diamond" (Challenge) Goal

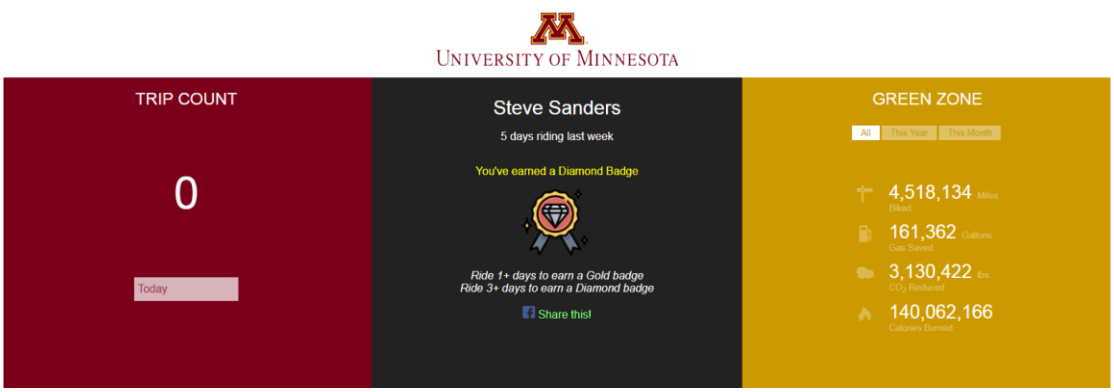


Figure 7-11: Dashboard with Challenge Badge Earned

No Badge Earned

Figure 7-12: Dashboard with Badge Unearned

No Badge Condition

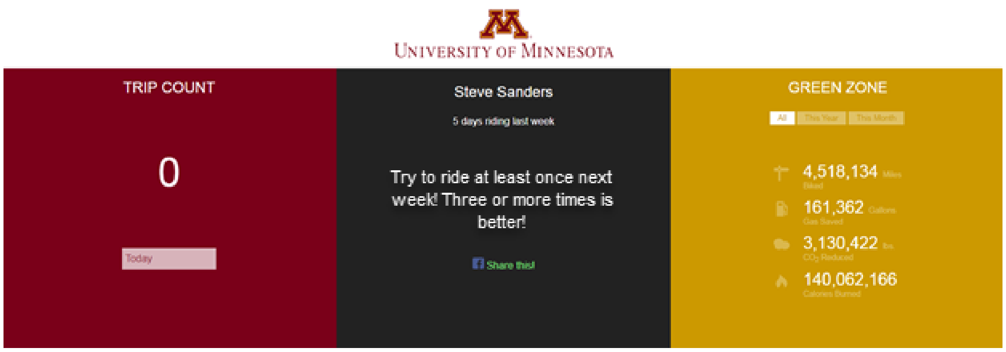


Figure 7-13: Dashboard with No Badge and Encouragement

Fifth (no encouragement) condition

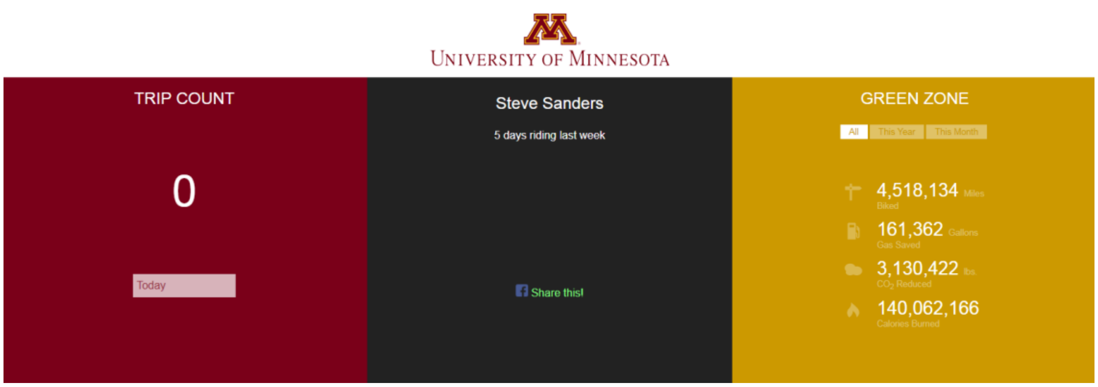


Figure 7-14: Dashboard with No Badge and No Encouragement

Appendix 3.2: Sample Emails

Emails are constructed using the same scripts which construct the dashboard panels with two changes: first, the colors are changed (to Minnesota colors) since the background is white rather than multi-colored. Furthermore, the font is changed due to security restrictions on emails.

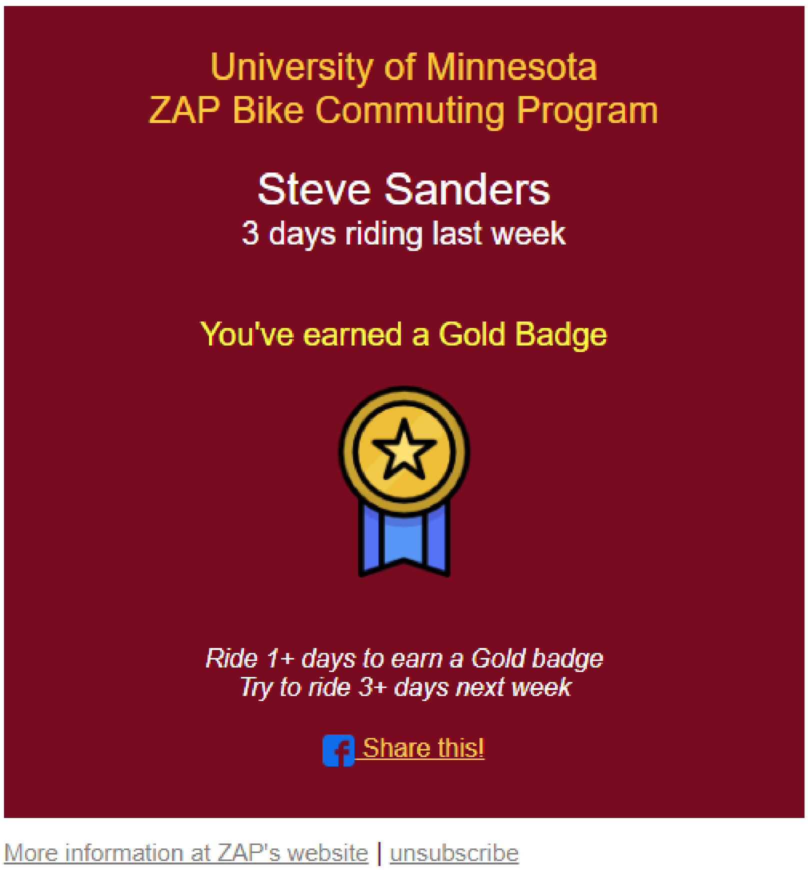


Figure 7-15: Sample Email with Participation Badge

Appendix 4.1: Dashboard Images

No Badge Earned

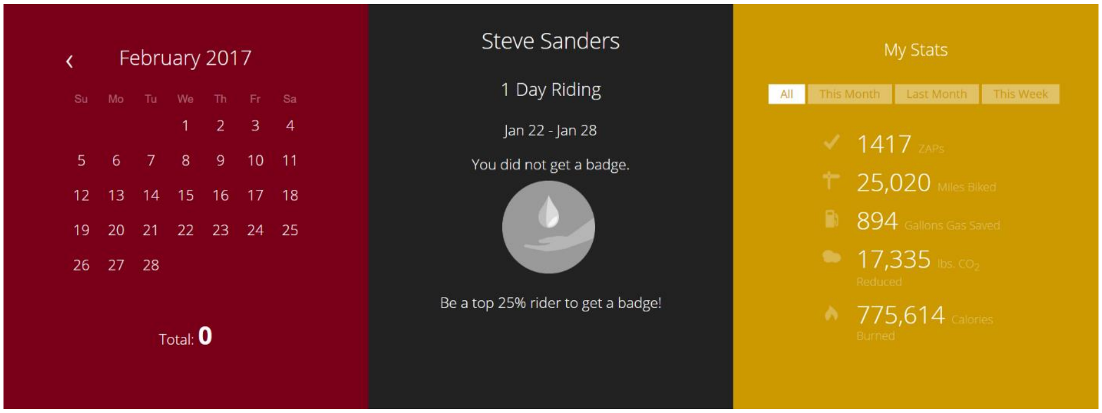


Figure 7-16: Prosocial Badge Unearned with Competitive Condition

Badge Earned

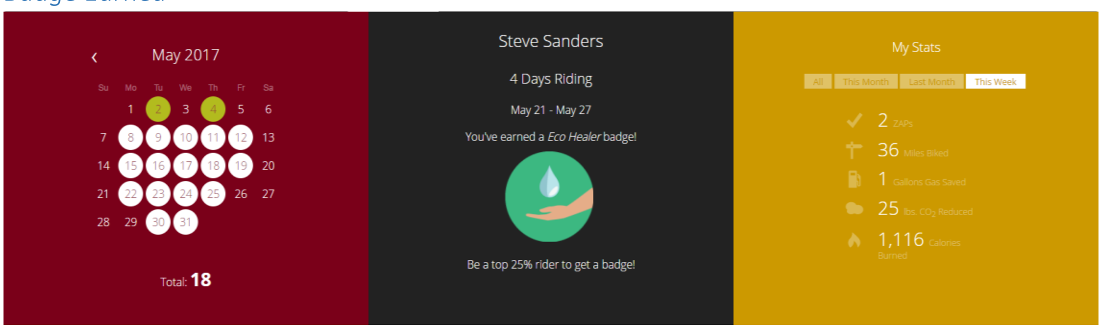


Figure 7-17: Prosocial Badge Earned

Appendix 4.2: Email Images

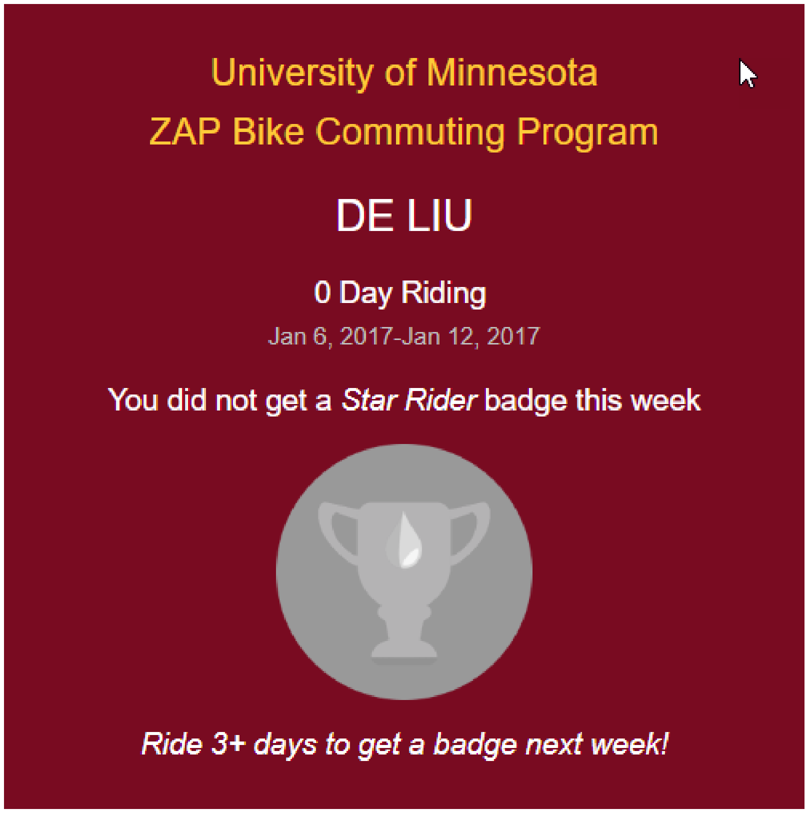


Figure 7-18: Sample Email with Achievement Badge Unearned

Appendix 4.3: Badge Designs

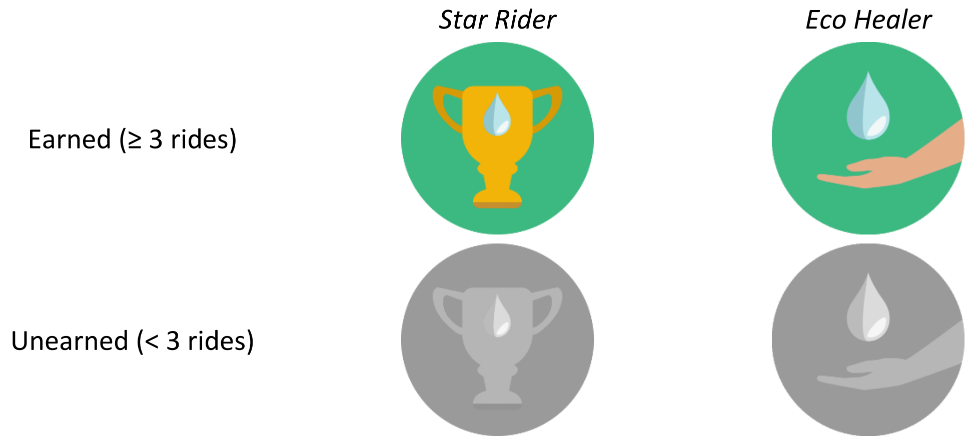


Figure 7-19: Badge Designs, Earned and Unearned