Can Gamification Motivate Voluntary Contributions? The Case of StackOverflow Q&A Community

Huseyin Cavusoglu
University of Texas at Dallas
800 W. Campbell Road
Richardson, TX 75080 USA
huseyin@utdallas.edu

Zhuolun Li
National University of Singapore
Information Systems Research Lab
Singapore, 117543, Singapore
zhuolun@comp.nus.edu.sg

Ke-Wei Huang
National University of Singapore
COM2, 04-18
Singapore, 117543, Singapore
huangkw@comp.nus.edu.sg

CSCW '15 Companion, Mar 14-18 2015, Vancouver, BC, Canada
ACM 978-1-4503-2946 0/15/03.
http://dx.doi.org/10.1145/2685553.2698999

Abstract
Online communities heavily rely on voluntary participation and continued engagement from users because these sites can flourish only if there are meaningful contributions from community members. Gamifying the underlying incentive mechanism can be a solution to elicit and sustain the desired user behavior. In this paper, we develop a theory of gamification and study the impact of a hierarchical badges system, a reward mechanism based on gamification principles, on user participation and engagement at Stack Overflow Q&A site. Specifically, we assess the extent to which users are incentivized by earned badges in their contributions to the answering activity. Our initial results present strong empirical evidence that confirms the value of the badges and the effectiveness of gamification in stimulating voluntary participation.

Author Keywords
Gamification; online communities; Q&A sites, voluntary contribution; badges; motivation theories.

Introduction
It is a challenging proposition for online communities to incentivize users to have active engagement in the community since users participate without getting

monetary rewards. Among the myriads of online communities, Stack Overflow question-and-answer (Q&A) site has a special place because of its unique motivation features. At the heart of its incentive scheme lies a reputation system that is based on peer approval, balancing the quality of contribution over the quantity of contribution. Users of this community assess the value of each other's contributions. Points, badges, and levels for quality participation are the symbols of achievements. Stack Overflow leveraged the power of the crowd for quality contributions by adopting the principles of gamification. *Gamification* refers to the use of game mechanisms and techniques in a non-game context to motivate and engage people across various activities [3]. In a nutshell, gamification acknowledges users' achievements, encourages them as they progress through levels, and eventually gets users emotionally engaged in the desired behavior.

Despite the adoption of various game design elements in online communities for engagement, there is limited research that theoretically argues and empirically quantifies the significance of gamification for voluntary activities. It is not clear if gamification is just an overhyped technology [4], or it has a potential to overhaul the old engagement models. This paper attempts to fill this gap by examining the value of a hierarchical badges system on user activities in an online Q&A community, namely Stack Overflow.

Theoretical Development

We conceptualize gamification through the lenses of cognitive evaluation theory and organismic integration theory [1,2]. Specifically, we argue Stack Overflow users contribute to the community because badges enhance motivations and support continued engagement. In other words, badges awarded serve as external events that impact intrinsic motivation (within the framework of cognitive evaluation theory), and promote the internalization process (within the framework of organismic integration theory) because received badges tap into psychological needs of users. First, earning a badge indicates *competence* in performing the activity as other community members (i.e., peers) had to confirm the value of your contribution (by voting up, making it favorite, choosing to read your contribution). An earned badge facilitates experience of effectance in a user's community activities as it delivers competence-relevant feedback. Second, receiving a badge promotes a sense of *autonomy*, which in turn facilitates the progress toward internalization. That is, even if users' contributing behavior is initially driven by expectation of earning a badge, once the badge is rewarded, this event provokes a motivational move toward feeling a sense of pride after engaging in the same act. Third, earning a badge promotes the experience of reciprocal care and concerns for other community members. A badge fosters the feeling *connectedness* with the Q&A community as it signifies and reinforces being valuable to other site users. Therefore, an earned badge, which is a signal of relatedness, motivates the user to participate more in same activity to feel more attached to the community. Hence, gamified Stack Overflow site can elicit continued contributions because earned badges not only enhance the intrinsic motivations of users, but also facilitate the internalization of their extrinsic motivations.

Dataset

We employed the Stack Exchange APIs to crawl data from August 1st 2008 to January 1st 2012. We excluded

the observations in 2008 because user activities were relatively unstable due to the nascence of Stack Overflow. Our final sample period included three years of data from January 1st 2009 to January 1st 2012. We next excluded users who have never earned 200 reputations. Our final dataset included 46,571 users. We then gathered user-level panel datasets, including a detailed history of reputation, types of badges received, and types of activities engaged in with a time stamp.

Empirical Analysis

To investigate the effect of individual badges, we conduct our empirical analysis using (i) *t*-test, (ii) difference-in-differences(DID) method, and (iii) propensity scoring matching (PSM). We rely on these techniques not only to show the impacts of earned badges on participation but also to check the robustness of our results to exclude alternative explanations.

We currently consider the answering activity only. Our analysis is conducted as follows. First, we identify each date that a user gets a specific badge. For the user who earned the target badge, we aggregate the number of answers given by the user to other users' questions; 7 days before and 7 days after badge awarding date, denoted by $Contributions_{i0}$ and $Contributions_{i1}$, respectively. We next compare the number of activities before and after obtaining the badge using a *t*-test.

Our DID analysis is conducted as follows: for each specific badge event *i*, the dependent variable is the first-differenced, weekly number of activities to eliminate any user-specific, time invariant heterogeneity [6]. Formally, our dependent variable is $\Delta Contributions_i = Contributions_{i1} - Contributions_{i0}$. The estimation model is given by $\Delta Contributions_i = \beta_{0i} + \beta_{1i} * Badge_i + \beta_{2i} * Controls_i + u_i$, where $Badge_i$ is a dummy variable and equals to one when a user receives the target badge *i*. $Badge_i$ is zero for all other users. In this estimation, β_{1i} captures the average badge (treatment) effect.

We also use PSM method suggested by [5] as an alternative method for identifying the badge effect. In PSM, the control group is formed by matched observations that are most similar to the observations in the treatment group. Our control group is constructed as follows. For each user-badge pair in the treatment group, we find the ten most similar users who did not receive the target badge 7 days before and 7 days after the treated user received the target badge. We use the 10-nearest neighbor-matching algorithm with replacement.

Preliminary Results

We report estimation results in Tables 1. Our results based on *t*-test in column (a) suggest that the impacts of all answers badges are both positive and significant. Estimation results from the DID model in column (b) confirm that every answers badge has a significantly positive impact on the answering activity. Furthermore, the PSM results in columns (c) reaffirm that answers badges induce users to engage more in the answering activity. All these results state that users who earn badges for answering questions subsequently answer more questions after earning these badges.

It is interesting to note that even questions badges give a positive boost to the answering activity. The estimation results for questions badges are all positive and significant in columns (a) and (b). Although the

impacts are positive, the effects seem small relative to the effects of answers badges. This result implies that badges induce more engagement in a broader set of activities (not just for the activity for which the badge is designed), presenting evidence that users internalize the incentives provided by earned badges.

Future Research

We plan to augment the current findings in several dimensions. First, we will consider other contribution activities and the impact of badges on these activities. Second, since badges are hierarchical and categorized into three groups (bronze, silver and gold), we will analyze the average effect of a badge category. Our objective is not only to quantify the influence of badge categories but also to assess the effectiveness of the hierarchical badges system. Third, it is possible that the stimulus provided by earned badges may change as users evolve from a novice to a senior community member. To examine this issue, we will estimate the model with separated time phases.

References

[1] Deci, E.L. *Intrinsic Motivation*. New York: Plenum. 1975.

[2] Deci, E.L., and Ryan, R.M. *Intrinsic Motivation and Self-Determinaton in Human Behaviour*. New York: Plenum. 1985.

[3] Deterding, S., Khaled, R., Nacke, L.E., and Dixon, D. "Gamification: Toward a Definition," In *CHI Gamification Workshop*, ACM Press (2011)

[4] LeHong, H., and Fenn, J. Hype Cycle for Emerging Technologies, Gartner Research, 2012.

[5] Rosenbaum, P.R., and Rubin, D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* 70,2 (1983) 41-55.

[6] Wooldridge, J.M. 2009. *Introductory Econometrics: A Modern Approach*. South-Western Pub.

Badge Name	***(a) T-test***	***(b) DID***	***(c) PSM***
Popular Question	*0.056*** (0.003)*	*0.077*** (0.004)*	*0.019*** (0.006)*
Notable Question	*0.042*** (0.005)*	*0.065*** (0.007)*	*-0.007 (0.009)*
Famous Question	*0.044*** (0.012)*	*0.068*** (0.017)*	*0.000 (0.023)*
Nice Question	*0.227*** (0.007)*	*0.325*** (0.010)*	*0.279*** (0.022)*
Good Question	*0.170*** (0.014)*	*0.255*** (0.020)*	*0.177*** (0.046)*
Great Question	*0.287*** (0.045)*	*0.378*** (0.063)*	*0.112 (0.177)*
Favorite Question	*0.231*** (0.023)*	*0.346*** (0.032)*	*0.223*** (0.076)*
Stellar Question	*0.407*** (0.063)*	*0.540*** (0.088)*	*0.099 (0.212)*
Nice Answer	*1.426*** (0.004)*	*2.319*** (0.006)*	*1.480*** (0.030)*
Good Answer	*0.828*** (0.009)*	*1.424*** (0.013)*	*0.570*** (0.053)*
Great Answer	*0.779*** (0.028)*	*1.213*** (0.038)*	*0.431*** (0.160)*
Revival	*0.202*** (0.011)*	*0.396*** (0.015)*	*0.276*** (0.037)*
Necromancer	*0.216*** (0.010)*	*0.356*** (0.014)*	*0.231*** (0.031)*

Table 1. The impact of badges on the answering activity. Robust standard errors are in parentheses: *** p<0.01, ** p<0.05, * p<0.1.