BEHAVIORAL MECHANISMS PROMPTED BY VIRTUAL REWARDS: THE SMALL-AREA HYPOTHESIS

Completed Research

Mutter, Tobias, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany, tobias.mutter@wiwi.uni-paderborn.de

Kundisch, Dennis, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany, dennis.kundisch@wiwi.uni-paderborn.de

Abstract

In a natural experiment on a popular German Question & Answer community we investigate the applicability of the small-area hypothesis to the activation of user contributions through virtual rewards in the form of badges. Koo and Fishbach's small-area hypothesis states that individuals in pursuit of a goal exhibit stronger motivation when they focus on whichever is smaller in size: the area of their completed actions or their remaining actions needed to reach a goal (e.g., focusing on 10% of completed actions is more motivating than on the 90% remaining). This has direct implications for the optimal design of virtual reward systems and especially for the framing of progress towards virtual rewards, which represent goals to users. Consistent with theoretical predictions, we find that the small-area effect activates online user contribution behavior. Our findings thus provide empirical evidence for the influence of the framing of progress towards virtual rewards on user behavior.

Keywords: Small-Area Hypothesis, Gamification, Virtual Rewards

1 Introduction

Over the last few years, gamification has experienced a rise in popularity and turned into a trending topic among practitioners and academics (e.g., Gartner 2011, Blohm & Leimeister 2013, Hamari et al. 2014). Gamification refers to the application of game design elements in a non-gaming context (Deterding et al. 2011), and is used by all manner of organizations for a variety of purposes: to improve user engagement, to motivate employees, to facilitate innovations, to promote personal development, to improve learning, and to encourage people to make healthy choices (e.g., Kumar 2013, Penenberg 2013, Burke 2014). Popular game elements include badges, points, levels, or leaderboards (Hamari et al. 2014). The popular question and answer site StackOverflow, for example, uses badges dubbed 'Guru' and 'Illuminator' to activate its members.

Gamification has its critics too. Burke, for example, suggested that gamification raises too many unrealistic expectations as it 'will move through the hype cycle from the peak of inflated expectations into the trough of disillusionment' (Burke 2013). In the same vein, Gartner (2012) predicts that '80 percent of current gamified applications will fail to meet business objectives primarily because of poor design'. Rather than dismissing the potential of gamification altogether, these criticisms open up the field to researchers. In particular, more research is needed to better understand the behavioral mechanisms associated with gamification. Such insights would enable gamification designers to integrate game elements into applications more successfully.

While research suggests that gamification can exert a positive effect on user motivation and engagement its impact depends on both the context and the precise manner in which game elements are implemented (e.g., Hamari et al. 2014). With our research we want to contribute to understanding the key drivers behind the effectiveness of gamification by analyzing the so-called *small-area hypothesis* in the context of online-communities. The small-area hypothesis states that individuals in pursuit of a goal exhibit stronger motivation when they focus on whichever is smaller in size: the share of completed actions or the share of actions still needed to reach a goal (Koo and Fishbach 2012). In other words, the framing of the recorded progress affects motivation. In practical terms, users who are in the early stages of goal-pursuit show greater motivation when presented with their accumulated progress (e.g. 10% achieved) rather than with their remaining progress (e.g., 90% remaining), whereas with greater proximity to the goal, it is more effective to focus users on their remaining progress (e.g. 10%) rather than on their accumulated progress (e.g., 90%).

The small-area hypothesis has already been documented in the context of customer loyalty programs (Koo & Fishbach 2012). However, given the substantive differences between loyalty programs and non-monetary virtual reward systems, it is by no means evident whether this finding can be extended from one to the other. Leaving aside the absence of monetary incentives or quasi-monetary benefits (e.g., lounge access or priority booking at frequent flyer programs), another main difference is that customer loyalty programs aim to influence individual decision making, notably buying behavior, while virtual reward systems in the context of online communities are designed to address motivational phenomena such as user effort. By answering the following research question, we investigate the generalizability of the small-area hypothesis to those aspects: *Does the small-area effect activate the contribution behavior of users in online communities?*

If the small-area effect can be shown to activate the contribution behavior of users, this would have important practical implications for the design of virtual reward systems. It would change the way designers of such systems should consider framing the distance or proximity towards a virtual reward, as well as help determine the optimal number of goals in a virtual reward system.

To address our research question we exploit a natural experiment by using a unique and rich dataset provided by a German Question & Answer (Q&A) community. This exclusive dataset includes detailed information about all user activity on the platform between February 2006 and April 2008. To activate its members, the platform has set up a virtual reward system. On performing certain, selected, activities, users are rewarded with points and by accumulating these points they can earn a series of badges. Thus, in our research setting goals are represented by badges. The natural experiment took place in February 2007, in the middle of our observation period, when the operator of the platform fundamentally restructured the virtual reward system. As a consequence users were exogenously set back from their next goal and the average distance towards the next badge increased. This natural experiment provides a unique research setting for the identification of the small-area effect. In our empirical analysis, we compare the contribution behavior of 650 users in the seven days before and after the event. We find that the users who were set right back to the beginning increase their post-event contribution levels, whereas users who were set back only half-way decrease their contribution levels. Since in both situations progress towards the next badge is framed in terms of accumulated actions we are able to explain this seemingly contradictory behavior with the small-area effect.

With this paper we make novel and significant contributions to research in two ways: (1) we contribute to the literature of gamification by providing empirical evidence that the framing of the progress towards virtual rewards affects user contribution levels; (2) we contribute to the research on the small-area hypothesis by being the first to provide empirical evidence of this effect in the presence of goals in form of non-monetary rewards, and by showing that the small-area effect also applies to motivational phenomena such as user effort.

2 Theoretical Background

One important finding from research on motivation is that persistence increases with proximity towards a goal's end state (Koo & Fishbach 2012). Research explains this phenomenon with the goal-gradient hypothesis (e.g., Hull 1932, Kivetz et al. 2006, Mutter & Kundisch 2014b). Kivetz et al. (2006). For example, in a field study conducted at a university café in which participating customers have to buy ten cups of coffee to get one for free, researchers found that participants purchase coffee more frequently the closer they get to the reward. A widespread explanation for this phenomenon is based on the perceived contribution of each consecutive action towards goal achievement which increases with proximity towards the goal's end state (Brendl & Higgins 1996, Förster et al. 1998). For example, buying the first of ten cups of coffee at the café reduces the distance to the goal by 10% (1 out of 10 outstanding cups), whereas purchasing the last cup reduces the distance by 100% (1 out of 1 outstanding cups).

Based on the view that the perceived impact of actions affects the motivation to perform the action, Koo & Fishbach (2012) propose the small-area hypothesis. The small-area hypothesis states that apart from the actual level of progress, motivation is also affected 'by the perception that the action has greater impact because the person is comparing it to a smaller set of other actions (e.g., stronger motivation for 20% completed vs. 80% remaining)' (Koo & Fishbach 2012, p. 507). This implies that motivation can be positively affected by being either far from or close to goal completion, because in both situations people are able to focus on whichever is the smaller area and hence, the one in which their action is perceived to have the greater impact (Bonezzi et al. 2011). However, this does not apply to the mid-point in a goal pursuit, regardless of how progress is framed. 'The small-area effect is orthogonal to the goal-gradient effect, such that both proximity to goal attainment and attention to small areas independently increase the perceived impact of an action and thereby increase motivation' (Koo & Fishbach 2012, p. 494).

In the field of marketing research, Koo & Fishbach (2012) provide empirical evidence for the small-area and the goal-gradient hypothesis in the context of customer reward programs. Their findings are consistent with the results from Bonezzi et al. (2011) in the field of psychology, who present evidence for a non-monotonic motivational pattern which consists of the classical increasing goal-gradient with proximity to the goal and a decreasing goal-gradient from the early stages of goal-pursuit. We contribute to this literature by empirically testing whether the small-area hypothesis also applies to virtual rewards system with non-monetary incentives and to motivational phenomena such as user effort.

3 Research Environment[1]

The website at the center of our analysis was launched in January 2006 and has requested to stay anonymous. The platform offers registered and non-registered users the opportunity to ask questions to the community on everyday topics (e.g., beauty, computers, gardening). In other words, the platform deals exclusively with leisure rather than labor-market related topics. All registered users automatically participate in the virtual reward system of the platform. For almost all of the activities performed, registered users receive an incentive in the form of *status points.* Each time users earn status points, their total number of status points increases. Users need to accumulate a predetermined total number of status points to earn badges. In Table 1, we present a list of the *main activities* and the corresponding status point scheme.

Main Activities	Status Points per Activity
Answering Questions	0 - 25
Asking Questions	0 - 4
Adding Friends	5 - 20
Adding & Copying Links	1 - 2

Table 1: Status Point Scheme (Before the Event)

[1] Two related papers by Mutter & Kundisch (2014a, 2014b) are drawing on the same research environment. Despite some overlap in the underlying dataset, the related studies differ in their scope, each addressing independent research questions.

The core activity on the platform is *answering questions*. Depending on the quality of their answer, users can earn between 0 and 25 status points for a given answer. The quality of the answer is rated by both the questioner and by other members of the community, but only the questioner can tag an answer as *top* answer whereas the members of the community can tag it as *helpful*. Apart from the activity *answering questions*, registered users can also get status points by *asking questions* to the community. If a question receives at least one answer or is rated as a *helpful* question by at least one other user, the questioner receives between 1 and 4 status points. No status points are earned, however, if the question remains unanswered. Registered users also have the opportunity to *add friends* to their network of friends. If a friend request is accepted by another user, both users earn a certain amount of status points. Furthermore, each user has a personal link catalogue. Whenever a user *adds a new link* to the catalogue or *copies a link* from another user, she earns status points.

In Table 2, we provide a detailed list of all the available badges and the total number of status points required for each badge. The badge ‘Bachelor’ (‘Master’), for example, requires an accumulation of at least 120 (720) status points. By earning an average of 4 status points per answer users would have to answer more than 30 (180) questions to earn this badge. Thus before users reach the ‘Master’ status they have to earn the ‘Bachelor’ badge.

Label of Badge	Required Status Points	Label of Badge	Required Status Points
Student	0	Archimedes	4,790
Bachelor	120	Ts’ai Lun	4,890
Master	720	Johannes Gutenberg	4,990
Research Assistant	1,130	Alexander G. Bell	5,090
Doctor	1,640	Gottfried W. Leibniz	5,190
Assistant Professor	2,250	Max Planck	5,290
Professor	3,050	Johannes Kepler	5,390
Nobel Laureates	3,780	Leonardo da Vinci	5,490

Table 2: List of Badges (Before Event)

The list with the badges and the required status points for each badge are publicly available on the platform. The badge and the total number of earned status points are displayed in each user’s personal profile. Both pieces of information are also publicly visible to other platform users or guests whenever a user poses or answers a question.

On this platform the level of progress towards the next badge is framed in terms of completed actions because users’ total number of status points is represented as an increasing number. It is important to note that the total number of status points is not reset to zero after users have earned a badge. This means that the small-area effect can activate user contribution behavior only shortly after users register on the platform, because only then do they possess a small total number of status points. However, it is more challenging to isolate the impact of the small-area effect directly after their registration from observational data alone, because there might be other factors at play that could affect user behavior. For example, users might be more passive in the earlier phases of their membership until they get to know the community better before starting to focus on goal attainment and adding their own contributions. Fortunately, a natural experiment that took place on the platform allows us to isolate the impact of the small-area effect.

4 Natural Experiment

In February 2007, the operator of the Q&A community fundamentally restructured the virtual reward system. According to the operator, the objective of the restructuring was to simplify and enhance the reward system. The provider changed the status point scheme for the activities on the platform, retrospectively recalculating the total number of status points of each user and modified the badge system. As a result of this restructuration, the number of status points that could be earned for certain activities listed in Table 3 were either *reduced* or *abolished*. These activities included *adding* and *copying links* and *adding friends*. The activities *asking* and *answering questions* were not affected by the restructuring. The new status point scheme is illustrated in Table 3.

Main Activities	Status Points per Activity		Status Points *reduced* or *abolished*?
	Before Event	*After* Event	
Answering Questions	0 - 25	0 - 25	(unchanged)
Asking Questions	0 - 4	0 - 4	(unchanged)
Adding Friends	5 - 20	0	✓
Adding & Copying Links	1 - 2	1	✓

Table 3: Status Point Scheme (After the Event)

In addition, the community provider *recalculated* the total number of status points that each user had earned since the first day of registration, based on the new point scheme. For example, by adding a new friend to their network users were rewarded with up to 20 status points before restructuration but none at all after the event – the reward for this activity had been abolished. Not only this, but if a user had earned 40 status points by adding new friends before the event, she lost these 40 status points after the event.

The new badge system is illustrated in Table 4. The provider added two new badges, changed the labels of the badges between 'James Watt' and 'Leonardo da Vinci' (see Table 2), and increased the number of required status points for each badge. The labels and the order of the badges from 'Student' to 'Nobel Laureates' and for the badge 'Albert Einstein' stayed the same. Users who held a badge between 'Student' and 'Nobel Laureates' *before* the event could compare their new position in the badge system based on the label of the new badge. Subsequently, these users could assess precisely how many badges they had lost. For example, a user with 200 status points held the badge 'Bachelor' *before* the event, while *after* the event, and holding the total number of status points constant, this user now holds the badge 'Beginner' and thus lost two badges.

Label of Badge after Event	Required Status Points after Event	Label of Badge as before Event?	Label of Badge after Event	Required Status Points after Event	Label of Badge as before Event?
Beginner	0	–	Robert Koch	8,240	–
Student	210	✓	Immanuel Kant	8,740	–
Bachelor	530	✓	Archimedes	9,240	–
Master	1,030	✓	Max Planck	9,740	–
Research Assistant	1,630	✓	Isaac Newton	10,240	–
Doctor	2,430	✓	T. A. Edison	10,740	–
Assistant Professor	3,330	✓	Pythagoras	11,240	–
Professor	4,240	✓	Galileo Galilei	11,740	–
Nobel Laureates	5,240	✓	Leonardo da Vinci	12,240	–
Albert Schweitzer	7,740	–	Albert Einstein	>12,740	✓

Table 4: List of Badges (After the Event)

The plan to restructure the virtual reward system was repeatedly announced prior to the implementation. The first announcement was made 5 months before the event. However, it is important for the following analysis that the specific modifications of the badge system - the recalculation and the deduction of status points - were not known to users in advance and had taken them by surprise.

As a consequence of the restructuring users were exogenously set back from their goal and the average distance towards their next badge was increased. This enables us to focus our analysis on two groups. The first comprises users who were set back to the beginning (and hence lost almost all of their status points) and the second, those who were only set back half-way towards earning the next badge (and hence lost fewer status points). As the positioning of the users after the event was determined exogenously, we have the opportunity to properly identify the small-area effect.

5 Hypothesis Development

According to the small-area hypothesis, we would expect to see an increase in the contribution levels of users who, as a result of the event, were set back to the beginning (with status points close to zero). This is because progress towards the next badge is framed in terms of completed actions. So when these users compare their recently earned status points to the lower (post-event) cumulative total of, say 10, compared with a pre-event total of 100, their post-event contribution is perceived as more effective (e.g. 4 points from one action added to 10, compared with 4 points added to 100, with next badge requiring 200 points). However, the impact of the small-area effect decreases as users accumulate status points. Thus, we would expect to see the post-event contributions of users who are placed around half-way towards the next badge to be only slightly - but positively - affected by the small-area effect. Therefore, we derive the following research hypothesis:

HYPOTHESIS: The online community users who are set back to the beginning are activated by the small-area effect and therefore increase their post-event contributions compared with users who are set back only half-way towards their next badge.

6 Dataset, Sample & Descriptive Statistics

6.1 Dataset

We are fortunate in having a unique dataset at our disposal which allows us to analyze this natural experiment provided by the community's operator. The whole dataset covers all user activities on the platform between February 2006 and May 2008. The number of newly registered users was 12,901 in 2006, 54,404 in 2007, and 25,909 up to the end of April 2008. During the observation period, we observe how these users collect 14,132,466 status points on the platform and, in the process, earn badges. To earn status points, users replied to 1,000,542 posted questions with 2,996,446 answers, built 32,696 friendships with other users, and added 87,872 links to the link catalogue of the platform. Our data is at the level of each individual user. Thus, we know exactly when a user registers on the platform, when and how often she performs a certain activity, when and how many status points she earns for her actions, and when she earns a badge. This allows us to establish a detailed profile for each user based on her activity history on the platform.

6.2 Sample

For our empirical analysis we select the 650 users who hold the badge 'Student' on the day prior to the event and who, at the time of the event, were still actively participating. We regard users as inactive if they permanently stopped performing any of the platform's activities. All users in our sample lost one badge and hold the badge 'Beginner' after the restructuring. In addition, these users lost status points and thereby were exogenously set back to an interval ranging from the beginning to half-way towards the next badge after the event. We choose this group of users because all users in this group receive the same treatment except for the positioning towards the next badge. In our empirical analysis, we compare the user contribution behavior of these users in the seven days before and after the event.

This leaves us with an unbalanced panel of 650 users and 8,650 observations on a daily level over a period of 14 days.

6.3 Descriptive Statistics

6.3.1 Activity History of Users

In Table 5 we present a short summary of the activity history for the 650 users in our sample from the foundation of the platform up to the day of the event. At the time of restructuration, users are on average registered on the platform for 99.6 days (*Length of Membership*), while 50% of users are registered for 52 days or more. During the entire period of their membership users contributed an average of 4 answers (*Sum of Answers*), asked 3.2 questions (*Sum of Questions*), had 0.2 friends (*Sum of Friends*), and added 0.4 links (*Sum of Links*).

Variables	Mean	Min	Q25	Median	Q75	Max	Sum
Length of Membership	99.6	1	14	52	155	392	–
Sum of Answers	4	0	0	1	6	47	2,612
Sum of Questions	3.2	0	0	1	3	49	2,089
Sum of Friends	0.2	0	0	0	0	3	126
Sum of Links	0.4	0	0	0	0	22	260

Table 5: Users' Activity History

6.3.2 Proximity to the Next Badge

In Figure 1, we present the distribution of users in our sample across five intervals which track the distance of users from the next badge *before* and *after* the event.

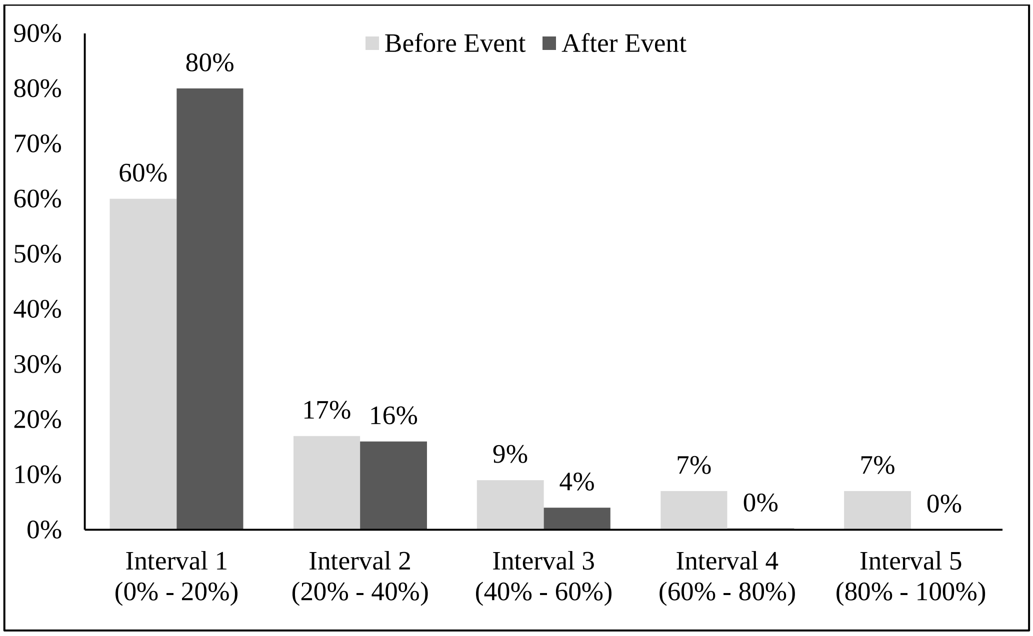

Figure 1: Proximity to the Next Badge

Each interval covers 20% of the required status points (e.g., *Interval 1* covers 0%-20% which is equal to the 0-24 status points before the event and 0-42 status points after the event). Before the event, 60% of users had earned less than 20% of the required points, 17% were positioned in *Interval 2*, and the remaining 23% of users were almost equally distributed across *Interval 3* - *Interval 5*. After the event, the distance towards the next badge increased substantially for those users. The proportion of users who possess less than 20% of the required points increased from 60% to 80%, and the remaining 20% are placed into *Interval 2* or *Interval 3*. After the event, no more users remain in *Interval 4* or *Interval 5*. We use this exogenous variation in the positioning of users in our empirical analysis to identify the small-area effect.

6.3.3 Quantity Measures

In Table 6 we illustrate the number of *Answers* and the number of *Main Activities* per user per day in the week *before* and *after* the event. The number of *Main Activities* represents the sum over the four main activities illustrated in Table 1. We provide mean, standard deviation, median, maximum value and the total sum for both variables. Naturally, we have a large number of zeros in our sample as we work with user activity data on a daily level. The average of *Answers* increases slightly from 0.13 per day before, to 0.14 after the event. The average for *Main Activities* increases also slightly from 0.33 to 0.34 activities per user per day.

	Before Event					*After* Event				
Variables	Mean	Std.	Median	Max	Sum	Mean	Std.	Median	Max	Sum
Answers	0.13	0.74	0.0	15	522	0.14	0.84	0.0	21	628
Main Activities	0.33	1.54	0.0	26	1,364	0.34	1.58	0.0	32	1,541

Table 6: Quantity of Users' Contributions

7 Empirical Analysis

7.1 Main Variables

We use the number of *Answers* per user per day to measure user contribution levels. In addition, we use the number of *Main Activities* as second quantity measure to rule out potential reallocation effects of effort (e.g., users might add fewer links while increasing the number of their answers). To test our research hypothesis, we create a dummy variable (*Small-Area Dummy*) which takes the value zero for users who are placed into *Interval 2* or *Interval 3* and one for users who are placed into *Interval 1* after the event. Finally, we create another dummy variable separating the days before and after the event (*Event Dummy*).

7.2 Model

We use a differences-in-differences (DD) approach to analyze the data from the natural experiment. With the DD framework we explicitly estimate how each group responds to the restructuring and how the response of each group differs. To consider the distribution properties of both quantity measures (i.e., only non-negative integer values and large number of zeros) we estimate a poisson model (Cameron & Trivedi 2013). The model is illustrated in equation (1):

$$Y_{it} = \alpha + \gamma D_S + \theta D_E + \rho(D_S * D_E) + \varepsilon_{it} \tag{1}$$

The variable Y_{it} represents the dependent variables. Each observation in the sample is identified exactly by the index it where i represents the individual and t the day in our observation period. The variable D_S is the *Small-Area Dummy*. The estimator for the coefficient γ reveals potential differences between the two groups in average activity levels before the event. D_E is the *Event Dummy* and the estimator for θ represents the difference in average activity levels of the first group between the seven days before and after the event. The coefficient ρ of the interaction term between the *Small-Area Dummy* and the *Event Dummy* reveals the difference between the differences in average activity levels for both groups. Hence, the estimator reveals the difference in how each group is affected differently by the restructuring. The variable ε_{it} is the error term. We cluster the standard errors on the user level to account for heteroscedasticity and autocorrelation in the data (Wooldridge 2010).

7.3 Identification

In the underlying research environment the level of progress towards the next badge is framed in terms of completed actions because the total number of users' status points is represented as an increasing number (see section 3). This implies that the small-area effect is the most pronounced when users are closer to zero status points and gradually weakens with an increasing number of points. This allows us to separate users into two groups, those who are set back to *Interval 1* and those who are set back to *Interval 2* or *Interval 3* (see section 6.3.2). Crucial to our analysis is the difference in each group's responses. Due to the small-area effect, users who are set back to *Interval 1* are expected to respond more positively to the event compared to users who are set back to *Interval 2* or *Interval 3*.

In equation (1), the estimator ρ for the interaction term between the *Small-Area Dummy* and the *Event Dummy* reveals how the responses between groups differ. There are two scenarios which can explain how the interaction term relates to the small-area effect. In the first, or *base case scenario,* both groups respond equally to the restructuring, were it not for the small-area effect. In this scenario the estimator for the *Event Dummy* θ is representative for both groups and the estimator for the interaction term ρ equals the small-area effect. In the second, or the *pessimistic scenario,* only users who are set back to *Interval 2* or *Interval 3* are negatively affected by the restructuring and the estimator for the *Event Dummy* θ is not representative for both groups. In this scenario the estimator for the interaction term ρ has to be substantially larger than the *Event Dummy* θ if it is able to identify the small-area effect. Otherwise the estimator for the interaction term ρ might only artificially mirror the estimator of the *Event Dummy* θ (e.g., θ≈-20% and ρ≈+20%).

In general, the *base case scenario* appears to be more likely than the *pessimistic scenario.* Both user groups are expected to be negatively affected by the event because the distance towards the next goal is increased after the event and thus the activating power of the goal-gradient effect is less pronounced (see section 2). However, as we cannot be absolutely certain of the presence of the base case scenario, we require the estimator for the interaction term ρ to be substantially larger in magnitude than the estimator for the *Event Dummy* θ, to enable us to identify the small-area effect in the subsequent analysis with confidence.

7.4 Results

In Table 7 we present the results of our empirical analysis. The first column shows the independent variables, the second column the results for the number of *Answers,* and the third column the number of *Main Activities*. For the dependent variable number of *Answers* all estimators are significant on a one percent level except for the *Event Dummy.*

Variables	*Answers*	*Main Activities*
Constant	-1.446** (0.230)	-0.684** (0.193)
Small-Area Dummy	-0.893** (0.269)	-0.577* (0.229)
Event Dummy	-0.348° (0.183)	-0.390* (0.159)
*Small-Area Dummy * Event Dummy*	0.642** (0.235)	0.560** (0.200)
Number of Users	650	650
Observations	8,650	8,650
-Ln Likelihood	-4,267	-8,943
Cluster Robust Standard Errors in Parentheses, ** p<0.01, * p<0.05, ° p<0.1		

Table 7: Empirical Results

The estimator for the *Event Dummy* is significant on a ten percent level. The estimator for the *Small-Area Dummy* is -0.893 or -60% and reveals that users who were set back to *Interval 1* were less active before the event than users who were set back to *Interval 2* and *Interval 3*. The estimator for the *Event Dummy* is -0.348 or -29%. This represents a decrease in the activity levels of users who were set back to *Interval 2* and *Interval 3*. The estimator for the interaction term between the *Small-Area Dummy* and the *Event Dummy* is 0.642 or 90%.

We find a similar pattern for the second measure of the contribution quantity. All estimators are significant on a one or five percent level. The estimator for the *Small-Area Dummy* is -0.577 or -44%, for the *Event Dummy* -0.390 or -32%, and the estimator for the interaction term is 0.560 or 75%.

7.5 Discussion

The negative estimators for the *Event Dummy* indicate that users who are set back to *Interval 2* or *Interval 3* decrease their activity levels after the restructuring. The positive estimators for the interaction term between the *Small-Area Dummy* and the *Event Dummy* indicate that users who are set back to *Interval 1* increase their activity levels after the event compared to users who are set back to *Interval 2* or *Interval 3*. Even more importantly, the estimators for the interaction term are substantially larger in size than the estimators for the *Event Dummie*s, which means that our results are valid for both the *base case scenario* and the *pessimistic scenario*. Thus, these results support the theoretical predictions which suggest that the activity levels of users who were set back to *Interval 1* are positively affected by the small-area effect. Hence, we derive the following result:

RESULT: The online community users who are set back to the beginning are activated by the small-area effect and substantially increase their post-event contribution levels compared with users who are set back only half-way towards the next badge.

This result provides support for our research hypothesis. If the framing of the progress towards the next badge had no impact on user activity levels, we would expect the activity levels of both groups to be negatively affected by the event. However, as the users who are set back to *Interval 1* are positively affected by the restructuring, we attribute this positive effect to the small-area effect.

7.6 Robustness Checks

Although we find support for our research hypothesis, we examine a number of robustness checks to demonstrate the robustness of our results.

7.6.1 Extended Model

We include the *Length of Membership* on the day before the event in absolute and squared terms in our model in equation (1) to account for negative effects of time (e.g., an increase in the probability to become inactive with increasing length of membership). The estimation results are illustrated in Table 8. The structure of the table is identical to Table 7. The estimator for the interaction term between the *Small-Area Dummy* and the *Event Dummy* is positive and significant for both dependent variables, that is, 0.502 or 65% for the number of *Answers*, and 0.437 or 55% for the number of *Main Activities*. Both estimators are lower compared to the estimators in Table 7. However, they are still reasonable in size and support the predictions from theory that the activity levels of users who were set back to *Interval 1* are positively affected by the small-area effect after the event.

Variables	*Answers*		*Main Activities*	
Constant	-0.181	(0.354)	0.456°	(0.271)
Small-Area Dummy	-1.319**	(0.286)	-0.959**	(0.228)
Event Dummy	-0.348°	(0.183)	-0.390*	(0.159)
*Small-Area Dummy * Event Dummy*	0.502*	(0.231)	0.437*	(0.199)
Length of Membership	-0.0230**	(0.005)	-0.0195**	(0.0036)
Length of Membership^2	0.00005**	(0.00001)	0.00004**	(0.00001)
Number of Users	650		650	
Observations	8,650		8,650	
-Ln Likelihood	-4,018		-8,420	
Cluster Robust Standard Errors in Parentheses, ** p<0.01, * p<0.05, ° p<0.1				

Table 8: Robustness Check I - Length of Membership

7.6.2 Adjusted Sample

We exclude the 394 users (60%) from our sample who were already positioned in *Interval 1* before the event (see Figure 1) and estimate the model in equation (1) for both dependent variables again. We adjust our sample to rule out that our findings are driven by users who were not set back after the event or who newly registered on the platform shortly before the event. The results are illustrated in Table 9 and structured in the same way as in the previous tables.

Variables	*Answers*		*Main Activities*	
Constant	-1.446**	(0.230)	-0.684**	(0.193)
Small-Area Dummy	-0.585°	(0.327)	-0.286	(0.277)
Event Dummy	-0.348°	(0.183)	-0.390*	(0.159)
*Small-Area Dummy * Event Dummy*	0.687*	(0.322)	0.579*	(0.269)
Number of Users	256		256	
Observations	3,572		3,572	
-Ln Likelihood	-2,230		-4,352	
Cluster Robust Standard Errors in Parentheses, ** p<0.01, * p<0.05, ° p<0.1				

Table 9: Robustness Check II - Adjusted Sample

For both dependent variables the estimator for the interaction term between the *Small-Area Dummy* and the *Event Dummy* is both positive and significant. The estimator for the number of *Answers* is 0.687 or 99%, and for the number of *Main Activities* it is 0.579 or 78%. Both estimators are higher compared to the estimators in our main model. This again supports our research hypothesis.

7.6.3 Additional Robustness Checks

(1) We include individual-specific fixed effects in the model in equation (1) to account for time constant user heterogeneity; (2) we adjust the size of *Interval 1* from 0%-20% to 0%-25% and 0%-30%, redefine the *Small-Area Dummy* variable, and run the model in equation (1) for each specification again; (3) we exclude the day of the event from our sample. Our main results remain qualitatively unchanged for each robustness check.

8 Conclusion

With this paper we enhance the understanding of the underlying behavioral mechanisms prompted by virtual rewards (badges) in online communities, drawing on the small-area hypothesis as an explanatory framework. We test the applicability of the small-area effect in a natural experiment which allows us to investigate whether the framing of the progress towards virtual rewards has any impact on user effort. We find an increase in user contribution levels in the core activity ‘answering questions’ when users are in the early stages of their goal pursuit and when their progress was framed in terms of accumulated actions (highlighting the 10% achieved instead of the 90% remaining). We further find evidence that the activating power of this effect weakens with increasing progress to the next badge. By providing empirical evidence for the small-area effect on user contribution levels in the context of virtual rewards, our results make a distinct contribution to the body of literature investigating gamification (e.g., Hamari 2014). In addition, we contribute to the research on the small-area hypothesis (Koo & Fishbach 2012) by extending its applicability to non-monetary goals and to motivational phenomena such as user effort.

Although we use a natural experiment to identify the small-area effect and thereby control for potential alternative explanations, we recognize that our results are not as robust as results from a randomized experiment. For example, it might be that some users increase their post-event activity levels because they are eager to regain their lost points. Although users in the treatment as well as in the control group lose points it might be that those users are unequally distributed across both groups. Future research could strengthen and refine our results by performing a randomized experiment with a two (progress: low vs. high) by two (framing: accumulated vs. remaining) between-subject design. Such an experiment would also provide the opportunity to investigate the interplay between the goal-gradient and the small-area effect in more detail. Another interesting approach for future research might be to analyze whether the framing of progress in large numbers is more effective in activating user contribution levels than framing in small numbers. Indeed, research suggests that the contribution of an action is perceived as higher when it is rewarded with a large number (e.g., 4,000 points) compared with a small number (e.g. 4 points) (Cantor & Kihlstrom 1987, Carver & Scheier 1998).

While the results from the Q&A community under study may not be directly transferable to other domains, our findings are nevertheless suggestive. Previous research in the domain of knowledge contribution has emphasized that user contribution behavior is influenced by both idealistic and altruistic factors (e.g., Kankanhalli et al. 2005, Jeppesen & Frederiksen 2006). We expect the small-area effect to be more pronounced in an environment where individuals are more extrinsically motivated and therefore more focused on virtual rewards and on their progress towards their reward goal. Thus, we have reason to believe that the activating power of the small-area effect could apply to various other domains including business and education.

Our results also have important managerial implications. Gamification designers should be aware that the framing of progress towards virtual rewards influences user effort. Our findings suggests that it would be more beneficial to frame progress in terms of accumulated actions in the beginning of goal pursuit up to a half-way point, and after this point is reached, to switch the framing to the number of actions remaining. For example, if a user needs 100 points to get a badge and has achieved 10% of the points, progress should be highlighted as ‘10% achieved’ and not as ‘90% remaining’. By contrast, when a user has earned 90% of the points, the progress should be presented as ‘10% remaining’ instead of ‘90% achieved’. The same reasoning also applies to any graphics illustrating progress (e.g., progress bar) which should highlight whichever is the smaller area of a user’s progress (accumulated progress or remaining progress). For example, if a user’s progress is represented by a solid blue line on a white background, the line should increase in length from 0 to 50%. When the midpoint is reached the colors of the progress bar should be inverted which means that the interval 0-50% is white and the interval 50-100% is blue. Beyond that point the solid blue line should decrease with increasing progress. This mechanism would ensure that a user focuses on whichever is smaller in size, regardless of whether this is the accumulated or the remaining progress. Finally, since the small-area effect appears to be effective in activating user contribution behavior shortly before and after users attain their goal, the existence of the small-area effect advocates a virtual reward system with multiple goals and medium achievement levels over a virtual reward system with fewer goals and higher achievement levels.

References

Blohm, I. and J. M. Leimeister (2013). "Gamification. Design of IT-Based Enhancing Services for Motivational Support and Behavioral Change." *Business & Information Systems Engineering* 5 (4), 275–278.

Bonezzi, A., C. M. Brendl and M. De Angelis (2011). "Stuck in the Middle: The Psychophysics of Goal Pursuit." *Psychological Science* 22 (5), 607–612.

Brendl, C. M. and E. T. Higgins (1996). "Principles of Judging Valence: What Makes Events Positive or Negative?" In: *Advances in Experimental Social Psychology*. Ed. by M. P. Zanna. SD: Academic Press, 95–160.

Burke, B. (2013). *The Gamification of Business*. URL: http://www.forbes.com/sites/gartnergroup/2013/01/21/the-gamification-of-business/ (visited on 11/28/2014).

Burke, B. (2014). *Gamify: How Gamification Motivates People to Do Extraordinary Things*. Bibliomotion, Inc.

Cameron, A. C. and P. K. Trivedi (2013). *Regression Analysis of Count Data*. Cambridge University Press.

Cantor, N. and J. F. Kihlstrom (1987). *Personality and Social Intelligence*. Englewood Cliffs, NJ: Prentice-Hall.

Carver, C. S. and M. F. Scheier (1998). *On the Self-Regulation of Behavior*. New York: Cambridge University Press.

Deterding, S., R. Khaled, L. E. Nacke and D. Dixon (2011). "Gamification: Toward a Definition." In: *Gamification Workshop Proceedings*. Vancouver: Canada.

Förster, J., E. T. Higgins and L. C. Idson (1998). "Approach and Avoidance Strength During Goal Attainment: Regulatory Focus and the 'Goal Looms Larger' Effect." *Journal of Personality and Social Psychology* 75 (5), 1115–1131.

Gartner (2011). *Gartner Says By 2015, More Than 50 Percent of Organizations That Manage Innovation Processes Will Gamify Those Processes*. URL: http://www.gartner.com/newsroom/id/1629214 (visited on 11/28/2014).

Gartner (2012). *Gartner Says by 2014, 80 Percent of Current Gamified Applications Will Fail to Meet Business Objectives Primarily Due to Poor Design*. URL: http://www.gartner.com/newsroom/id/2251015 (visited on 11/28/2014).

Hamari, J., J. Koivisto and H. Sarsa (2014). "Does Gamification Work? – A Literature Review of Empirical Studies on Gamification." In: *Proceedings of the 47th Hawaii International Conference on System Sciences*, 3025–3034

Hull, C. L. (1932). "The Goal-Gradient Hypothesis and Maze Learning." *Psychological Review* 39 (1), 25–43.

Jeppesen, L. B. and L. Frederiksen (2006). "Why Do Users Contribute to Firm-Hosted User Communities? The Case of Computer Controlled Music Instruments." *Organization Science* 17 (1), 45–63.

Kankanhalli, A., B. Tan and K. K. Wei (2005). "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation." *MIS Quarterly* 29 (1), 113–143.

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

Koo, M. and A. Fishbach (2012). "The Small-Area Hypothesis: Effects of Progress Monitoring on Goal Adherence." *Journal of Consumer Research* 39 (3), 493–509.

Kumar, J. (2013). "Gamification at work: designing engaging business software." In: *Proceedings of Second International Conference on Design, User Experience and Usability*. Las Vegas, NV, 528–537.

Mutter, T., and D. Kundisch (2014a). "Don't Take Away my Status! – Evidence from the Restructuring of a Virtual Reward System." *Computer Networks* (75:B), 477–490.

Mutter, T. and D. Kundisch (2014b). “Behavioral Mechanisms Prompted by Badges: The Goal-Gradient Hypothesis.” In: *Proceedings of the 2014 International Conference on Information Systems (ICIS 2014).* Auckland: New Zealand.

Penenberg, A. L. (2013). *Play at Work: How Games Inspire Breakthrough Thinking*. New York. USA: Penguin.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge.