Summer 6-15-2016

CHALLENGE ACCEPTED! - THE IMPACT OF GOAL ACHIEVEMENT ON SUBSEQUENT USER EFFORT AND THE IMPLICATIONS OF A GOAL'S DIFFICULTY

Tobias von Rechenberg
Polynomics AG, Olten, tobias.vonrechenberg@polynomics.ch

Dominik Gutt
Paderborn University, dominik.gutt@upb.de

CHALLENGE ACCEPTED! - THE IMPACT OF GOAL ACHIEVEMENT ON SUBSEQUENT USER EFFORT AND THE IMPLICATIONS OF A GOAL'S DIFFICULTY

Research

von Rechenberg, Tobias, Polynomics AG, Olten, CH, tobias.vonrechenberg@polynomics.ch

Gutt, Dominik,[1] Paderborn University, Paderborn, Germany, dominik.gutt@upb.de

Abstract

We empirically investigate the impact of successful goal achievement on future effort to attain the next goal in a recurring goal framework. We use data from a popular German Question & Answer community where goals are represented in the form of badges. In particular, our analysis of this data hinges on the fact, that in this Question & Answer community, badges in a hierarchical badge system are increasingly challenging to attain up to a certain badge. After this certain badge, the difficulty level suddenly drops and remains constant throughout up to the last badge in the hierarchy. Our findings indicate that after successful badge achievement users increase their subsequent effort to attain the next badge, but only as long as badges represent a challenge to the user. According to our analysis, we identify self-learning to be the key driver of this behavior.

Keywords: Goal Setting Theory, Recurring Goals, Self-Learning, Gamification, Badges

1 Introduction

"Success begat more success, which is the essence of positive feedback."

Information Rules (Shapiro and Varian 2013, p.174)

Gamification – the application of game design elements in a non-gaming context (Deterding et al., 2011) – has become one of the fastest growing business trends in recent years (Burke, 2013; Rauch, 2013). Its principles and techniques are used to motivate participation in various types of activities (Burke, 2014). For example, popular online communities like Wikipedia or StackOverflow use game design elements like badges, points, levels, or leaderboards to activate its users to contribute to the platforms' online activities (e.g., Anderson et al., 2013, Denny, 2013). Gartner (2011) predicts that by 2014 'more than 70 percent of Global 2000 organizations will have at least one gamified application'. However, at the same time, estimates suggest that '80 percent of current gamified applications will fail to meet business objectives primarily because of poor design' (Gartner, 2012). This highlights the necessity for a more systematic understanding of gamification to help developers to successfully integrate game elements into applications. Research suggests that gamification can have a positive effect on contribution behavior (e.g., Hamari et al., 2014), however, the question of why, when and how gamification works needs to be studied in more detail. Our research contributes to this knowledge by drawing on research on goal-setting and self-learning (Dzewaltowski et al., 1990, Ryan 1970, Drèze and Nunes, 2011). In particular, we answer the following research question:

Does successful goal achievement increase future effort to attain the next goal in a hierarchical goal system?

[1] This work was partially supported by the German Research Foundation (DFG) within the Collaborative Research Centre "On-the-Fly Computing" (SFB 901).

Most gamified applications provide virtual rewards in the form of levels or badges which represent goals to users (e.g., Hamari et al., 2014). On the one hand, research on the goal-gradient hypothesis indicates that users increase their contribution levels with proximity towards a goal and reduce their effort immediately after successful goal achievement (e.g., Kivetz et al., 2006, Mutter and Kundisch, 2014). This suggests that changes in user contribution levels are only temporary and driven purely by the user getting nearer to their goal. In contrast, research on goal-setting and self-learning indicates that successful goal achievement has a positive impact on future effort to achieve the next goal in a recurring goal framework as long as goals remain challenging (Drèze and Nunes, 2011). The iterative nature of this positive effect is of particular interest to both academics and practitioners in the gamification context.

To address our research question we are able to use 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 2007 and May 2008. Overall, we analyze the contribution behavior of 12,784 users over a time period of 462 days. To activate its members, the platform has set up a hierarchical badge system. On performing certain, selected, activities, users are rewarded with points and by accumulating these points they can earn 20 different badges. The level of difficulty increases stepwise for the first 10 badges which means that users need to collect an increasing number of points for each subsequent badge. At the 11th badge the level of difficulty drops to one fifth of the points needed for the preceding badge and remains constant at this level until the last badge in the incentive hierarchy. We use this break in the level of difficulty to identify the impact of badge achievement on subsequent effort levels to attain the next badge in this hierarchical badge system.

Our analysis reveals that users increase their effort levels after achieving a badge as long the level of difficulty keeps increasing, but hold their effort constant when the level of difficulty is being reduced. Thus, our results provide evidence for a positive effect of badge achievement on subsequent user contribution levels, through the channel of self-learning, but only as long as the achievement of badges continues to be challenging. 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 for the impact of successful badge achievement on user contribution levels in a hierarchical badge system; (2) we contribute to the research on goal-setting and self-learning by being the first to provide empirical evidence of this effect with goals in form of virtual rewards, and by showing that self-learning and self-efficacy also applies to motivational phenomena such as user effort.

2 Theoretical Background

Three strands of literature are relevant to our study. The first is related to research on goal setting, the second discusses the goal-gradient hypothesis, while the third introduces recurring goals and learning. These are presented in turn.

2.1 Goal Setting Research

Research on goal pursuit has a long tradition in psychology (Latham and Locke, 2007). Locke and Latham performed a comprehensive literature review to summarize the findings in the literature (Locke and Latham, 1990, 1991), revealing that specific and challenging goals lead to higher performance levels compared with easy or ‘do your best goals’ (Locke and Latham, 2002). As an explanation for this result, the literature suggested three key mechanisms that are also integral parts of Locke and Latham’s goal setting theory: goals (1) activate individuals to increase their effort; (2) lead to greater persistence; and (3) direct attention toward goal-relevant activities (e.g., Heath et al., 1999, Locke and Latham, 2002). In the context of our work, we focus especially on the increased user effort stimulated by goals and we focus on user persistence which is elicited through increasing goal difficulty.

2.2 Goal-Gradient Hypothesis

The goal-gradient hypothesis states that the effort to reach a goal increases with proximity towards a goal (e.g., Hull, 1932, Hull, 1934, Heilizer, 1977). Koo and Fishbach (2012) provide an overview of the different explanations for the goal-gradient hypothesis. For example, work on prospect theory uses the principle of diminishing sensitivity to explain that the marginal value of each action increases with proximity towards the goal which results in the willingness to bear higher costs for performance (Heath et al., 1999). Several recent studies have provided empirical evidence for the goal-gradient hypothesis based on humans (e.g., Kivetz et al., 2006, Mutter and Kundisch, 2014a). For example, in the context of online communities, Mutter and Kundisch (2014) show that users increase their effort levels with proximity towards goals in the form of virtual rewards.[2]

2.3 Recurring Goals and Self-Learning

Drèze and Nunes (2011) propose self-learning as a key driver for the impact of successful goal achievement on future effort to attain the next goal in a recurring goal setting. The theoretical reasoning for self-learning is based on self-efficacy. Self-efficacy refers to a person's belief about his or her ability and capacity to accomplish a task (Bandura, 1982, 1988). Previous research indicates that successful goal achievement can strengthen self-efficacy if a task is perceived as challenging in the sense of rendering the successful goal achievement uncertain (Dzewaltowski et al., 1990, Ryan, 1970). Building on this, Drèze and Nunes (2011, p. 270) state that 'if a task is challenging enough, each successful goal attainment should lead to a reassessment and, in turn, an increase in the base level of effort […]'. By extension, this mechanism implies that the positive effect of self-learning can occur multiple times as long as goals remain challenging. To sum it up, the answer to the question why people increase their effort in a recurring goal framework with increasing goal difficulty is, because they experience (*self-learn*) success upon goal achievement which leads to higher self-efficacy. In turn, higher self-efficacy is the driver behind the increased subsequent effort level. However success that comes too easily and which does not require persistence towards the goal is unlikely to lead to self-efficacy (Wood and Bandura, 1989). Therefore, only goals that are more challenging compared to the previous goal, are expected to lead to a higher base level of effort.

Concerning empirical evidence, Drèze and Nunes (2011) were also the first to analyze how successful goal achievement affects future effort to attain the same goal again in a recurring goal setting. They use data from a frequent-flier program to show empirically that customers who successfully earned the frequent flyer status within one year flew more frequently in the subsequent year compared with customers who failed to earn sufficient points to attain the frequent flyer status. In addition, they replicate their findings in a laboratory study and reveal that the positive impact of success on future effort is only prevalent when the goal is challenging. They identify self-learning as the key mechanism for explaining their results. With our research we extend the work from Drèze and Nunes (2011) to virtual reward systems which offer non-monetary benefits. In addition, whereas Drèze and Nunes focused their analysis mainly on decision making in the form of buying behavior, we investigate whether these results also apply to motivational phenomena such as user effort (i.e., contribution quantity of users). Due to the growing importance of websites that use gamification elements to foster user's contribution such as StackOverflow or Wikipedia, we consider this investigation a worthwhile contribution to theory that

[2] In a concurrent but independent working paper, Goes et al. (2016) analyse the impact of badges on user contribution behaviour in the context of an IT-related community. One of their findings is that users increase their contributions with decreasing distance to the goal. As Goes at et al. (2016) use data from an IT community on a weekly basis, first, our study differs in context, data granularity and theoretical scope. Second, while Goes et al. (2016) use the aforementioned results to investigate the overall impact of goals on performance, we integrate the results based on the goal-gradient hypothesis to analyse the incremental effect of goal achievement on the base level performance in a recurring goal framework.

does not, to the best of our knowledge, exist up to now. The next section introduces the research environment, followed by the explicit formulation of our research hypothesis.

3 Research Environment[3]

The website at the center of our analysis was launched in January 2006 and will remain anonymous at the owner's request. The platform offers registered and non-registered users the opportunity to ask questions to community members on everyday topics (e.g., beauty, computers, gardening). In other words, the platform deals exclusively with leisure-related topics, rather than labor-market related. All registered users automatically participate in the virtual reward system of the community. For almost all the activities they perform, registered users receive an incentive in the form of so-called status points. In Table 1, we present a list of the main activities and the corresponding status point scheme. Almost all (99%) status points are earned by users taking part in one of the two main activities, *asking* and *answering questions*. A few other activities (e.g., *inviting new members to the platform*) play only a very minor role, accounting for less than 1% of the total of accumulated status points.

Main Activities	Status Points per Activity	Average of Status Points Received	Ratio of Total Status Points
Answering Questions	0 – 25	4	76%
Asking Questions	0 – 4	3	23%

Table 1. Status Point Scheme

More specifically, an overall 76% of accumulated status points are earned with the activity *answering questions*. Depending on the quality of their answer, users can earn between 0 and 25 status points per 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* answer. On average, users earn 4 status points per answer.

Registered users can also earn status points by *asking questions* to the community. If a question receives at least one answer or is rated as *helpful* 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. On average, users earn 3 status points per question.

As they accumulate status points, users automatically move up in an ascending ranking system of 20 hierarchical badges. To earn a badge, users need to earn a predetermined number of status points, which varies from badge to badge. Table 2 provides a detailed list of available badges and the number of status points required. For example, the 'Master' badge requires an accumulation of at least 1,030 status points. Given the average of 4 status points per answer, users would need to answer more than 250 questions to earn this badge. The list with the badges and the required status points for each badge are also publicly available on the platform. The badge and the total number of earned status points are displayed in the user's personal profile. Both pieces of information are also publicly visible to other platform users or guests when posing or answering a question.

The hierarchical badge system on this platform represents a recurring goal framework: the level of difficulty increases constantly over the first 10 badges, then drops suddenly, and remains constantly low over the last 10 remaining badges. More specifically, the number of required status points increases from 210 for the badge 'Student' (*Badge 1*) to 2,500 for the badge 'Albert Schweitzer' (*Badge 10*), then it drops suddenly to 500 at the badge 'Robert Koch' (*Badge 11*) and from there on remains constant until the most valuable badge 'Albert Einstein' (*Badge 20*) is reached.

---

[3] Four related papers by Mutter and Kundisch (2014a, 2014b, 2015) and von Rechenberg et al. (2016) 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.

Nr.	Label of Badge	Required Status Points	Difference Status Points	Nr.	Label of Badge	Required Status Points	Difference Status Points
1	Beginner	0	0	11	Robert Koch	8,240	500
2	Student	210	210	12	Immanuel Kant	8,740	500
3	Bachelor	530	320	13	Archimedes	9,240	500
4	Master	1,030	500	14	Max Planck	9,740	500
5	Research Assistant	1,630	600	15	Isaac Newton	10,240	500
6	Doctor	2,430	800	16	T. A. Edison	10,740	500
7	Assistant Professor	3,330	900	17	Pythagoras	11,240	500
8	Professor	4,240	910	18	Galileo Galilei	11,740	500
9	Nobel Laureates	5,240	1,000	19	Leonardo da Vinci	12,240	500
10	Albert Schweitzer	7,740	2,500	20	Albert Einstein	>12,740	500

Table 2. List of Badges

4 Hypothesis Development

In line with previous research on the goal-gradient hypothesis (e.g., Kivetz et al., 2006, Mutter and Kundisch, 2014a), we would expect to see users increase their effort levels as they get closer to a badge: they always know how many points they need to achieve for the next badge and thus are able to trace their progress towards it. The pure goal-gradient effect on user effort is illustrated on the left side of Figure 1. Hence, we formulate our first hypothesis as follows:

HYPOTHESIS I: Users increase their effort the closer they get to a goal.

If, in addition, users experience self-learning after successfully achieving a badge, we would also expect them to increase their subsequent level of effort invested to reach the next badge, as depicted on the right side of Figure 1[4]. According to the theory of recurring goals and self-learning, the necessary condition for self-learning is that users perceive a badge as challenging. This would mean that, immediately after having accomplished the challenge of earning a badge their motivation carries them on to the next badge. Recall from 2.3 that motivation due to success only sets in, when the success does not come too easily, in other words, when the goal is more challenging than the previous goal. However, if a badge no longer presents a sufficient challenge, their motivation to continue reaching the next goal wavers. In other words, users are re-evaluating their self-efficacy after each badge. Also, the positive impact caused by self-learning which propels the user from one (challenging) badge achievement to the next, can occur repeatedly as long as the upcoming goals are perceived as challenging. On the right side of Figure 1 we illustrate the emerging activity pattern which is caused by the goal-gradient hypothesis and repeated self-learning. This leads to the next research hypothesis:

HYPOTHESIS II: After successful goal achievement, users increase their base level of effort when attempting to reach the next goal in a hierarchical goal system but only if achieving a goal is challenging.

[4] The abbreviations AP and AME in Figure 1 refer to the coefficients estimated in our model, which are introduced in our results section.

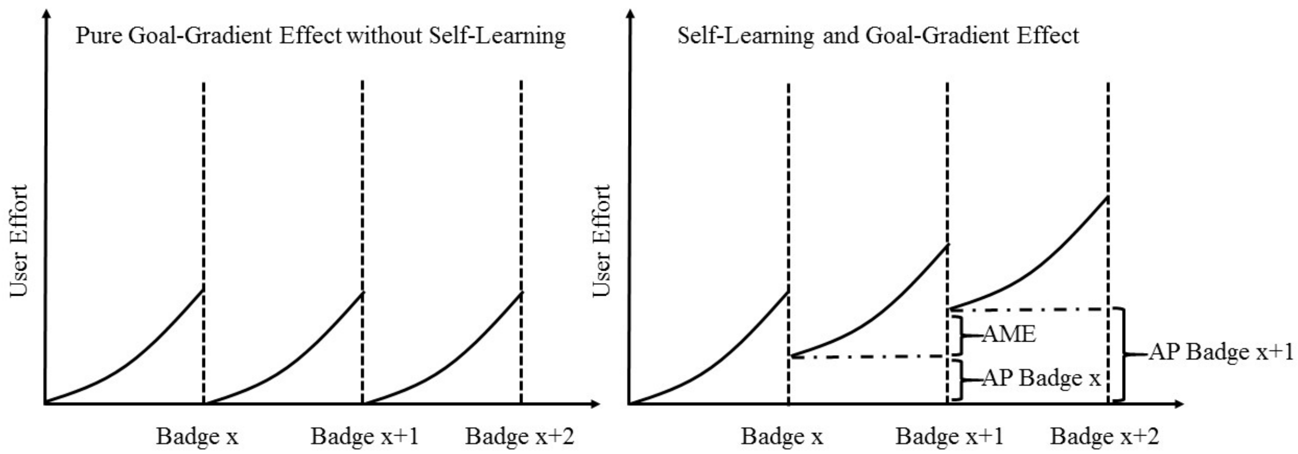

Figure 1. Self-Learning (Drèze and Nunes, 2011)

5 Dataset, Sample & Descriptive Statistics

5.1 Dataset

We are very fortunate in having a unique dataset at our disposal – kindly provided by the operator of this Q&A community – which allows us to analyze how successful badge achievement affects the subsequent base level of effort. The entire dataset covers all user activities on the platform between February 2007 and May 2008, i.e., an observation period of 462 days. During this observation period, 316,142 unregistered visitors posed a question to the community, and 73,017 new users registered on the platform. Our dataset enables us to observe that these users replied to 874,927 posted questions with 2,520,192 answers. Having data on the user level enables us to know exactly when a user registers on the platform, when and how often they perform a certain activity, when and how many status points they earn for their activities, and when they earn a badge.

5.2 Sample

We aggregate the activity data on a daily level to track how users adjust their base level of effort after successful badge achievement. We keep users in our sample (i.e., the corresponding observations), as long as they are active by participating in the recorded platform activities. We drop all observations where users hold the most valuable badge 'Albert Einstein' (*Badge 20*) because this badge represents the last goal in the recurring goal framework. This leaves us with an unbalanced panel of 12,784 users and 2,197,180 observations on a daily level over a period of 462 days.

5.3 Descriptive Statistics

5.3.1 Activity History of Users

Table 3 presents selected descriptive statistics for our sample. On average, our users are observed for 171.9 days (*Sum of Active Days*) before they become inactive and stop contributing to the platform. At the end of our observation period, users are registered for 195.9 days on average (*Length of Membership*). The maximum value of 885 days indicates that there are users in our sample who registered on the platform directly after the platform was launched in January 2006. For the duration of their membership users contribute an average of 169.5 answers each (*Sum of Answers*) and ask 33.2 questions (*Sum of Questions*). As can be seen from the distributions' quantiles, there is a strong heterogeneity in the history of user participation. For the main activities the median and mean values differ significantly. This indicates that the largest share of activities is performed by a small number of contributors.

Variables	Mean	Min	Q25	Median	Q75	Max	Sum
Sum of Active Days	171.9	1	45	145	279	462	2,197,180
Length of Membership	195.9	1	47	157	301	885	-
Sum of Answers	169.5	0	10	28	109.5	6,365	2,167,203
Sum of Questions	33.2	0	3	10	27	1,702	424,128

Table 3. Users' Activity History

5.3.2 Distribution of Badges

The users in our sample earn a total of 16,857 badges over the observation period. Table 4 illustrates the distribution of earned badges across the users in our sample. When they register on the platform users automatically receive the badge 'Beginner' (*Badge 1*), but from then on they need to collect more status points if they want to gain a more valuable badge. For the badge 'Student' (*Badge 2*), users need to earn 210 status points (see Table 2). We observe 5,345 users who collect sufficient status points to earn this badge. In general, the harder a badge is to obtain, the fewer users earn it.

Nr.	Label of Badge	Number of Promotions	Nr.	Label of Badge	Number of Promotions
1	Beginner	-	11	Robert Koch	247
2	Student	5,345	12	Immanuel Kant	224
3	Bachelor	3,315	13	Archimedes	208
4	Master	2,087	14	Max Planck	193
5	Research Assistant	1,487	15	Isaac Newton	183
6	Doctor	1,000	16	T. A. Edison	163
7	Assistant Professor	726	17	Pythagoras	161
8	Professor	539	18	Galileo Galilei	150
9	Nobel Laureates	422	19	Leonardo da Vinci	141
10	Albert Schweitzer	266	20	Albert Einstein	-

Table 4. Distribution of Badges

5.3.3 Quantity Measures

In Table 5, we provide mean, standard deviation, median, 95% quantile, 99% quantile, and maximum value for the number of *Answers*, the number of *Questions*, and the number of *Answers* & *Questions* per day on the user level. The number of *Answers* & *Questions* represents the sum over the number of *Answers* and the number of *Questions* per user per day. On average, users provide 0.99 answers and ask 0.19 questions per day. Naturally, we have a large number of zeros in our sample as we work with user activity data on a daily level.

Variables	Mean	Std.	Median	Q95	Q99	Max
Answers	0.99	4.49	0	5	21	218
Questions	0.19	1.04	0	1	4	254
Answers & Questions	1.18	4.88	0	7	23	254

Table 5. Quantity of Users' Contributions

6 Empirical Analysis

6.1 Main Variables

In our empirical analysis we use the number of *Answers* & *Questions* to measure user effort because both activities account for 99% of the overall acquired status points (see Table 1). Moreover, the activity asking questions (*Questions*) is more pronounced directly after users register on the platform while the activity answering questions (*Answers*) becomes more important with increasing length of membership. To investigate how badge achievement affects the subsequent base level of effort to attain the next badge we create a set of dummy variables which covers all badges on the platform (*Badge(1)* to *Badge(19)*). Each dummy variable measures the base level of effort associated with a particular badge. To account for the increase in user effort driven by the progress towards a badge, we include the relative number of earned status points towards the next badges (*Relative Progress*) in absolute and squared terms. This variable represents a continuous number between 0 and 1 and increases with proximity towards a badge. In addition, we interact this variable with the badge dummies to account for the differences in the level of difficulty of each badge (see Table 2). Finally, we include the variable *Length of Membership* in absolute and squared terms to account for user experience. We use both of these variables in absolute and squared terms to account explicitly for potential nonlinear effects.

6.2 Model

We estimate a Poisson model which is illustrated in equation (1).

$$y_{it} = \alpha + \sum_{\tau=2}^{19} \beta_\tau \times Badge(\tau) + \gamma_1 \times Relative\ Progress + \gamma_2 \times (Relative\ Progress)^2 + \sum_{\tau=2}^{19} \delta_\tau \times Badge(\tau) \times Relative\ Progress$$
$$+ \sum_{\tau=2}^{19} \rho_\tau \times Badge(\tau) \times (Relative\ Progress)^2 + \gamma_3 \times Lenght\ of\ Membership + \gamma_4 \times (Length\ of\ Membership)^2 + \varepsilon_{it} \quad (1)$$

Each observation in the sample is identified exactly with the index it where i represents the individual and t the day in our observation period. We include in the model all the variables described in the previous subsection. We use a Poisson model to account for the distribution properties of the dependent variable (i.e., only non-negative integer values and large number of zeros) and cluster robust standard errors to account for overdispersion and autocorrelation in the data (Wooldridge, 2010).

6.3 Identification

We use the estimators for the badge dummies (*(Badge (1)* to *Badge (19)*) to trace how the base level of effort changes as users move up in the hierarchical badge system. As described in section 3 and illustrated in Table 2, the badges become more and more challenging over the first 10 badges on the platform as the number of required status points for each badge increases from 210 to 2,500. At *Badge (11)* the level of difficulty suddenly drops to a mere 500 status points. According to theory, users are expected to increase their subsequent base level of effort after successful badge achievement when achieving a badge is challenging. Therefore, we use the rise in the level of difficulty over the first 10 badges and the sudden drop in the level of difficulty at the 11th badge to identify the impact of successful badge achievement on the base level of effort. If the base level of effort increases constantly from *Badge (1)* to *Badge (10)* and stops to increase at *Badge (11)* this would support the theoretical predictions. By exploiting this adjustment in the level of difficulty we can also rule out that our results are purely driven by unobserved user characteristics or user self-selection (e.g., intrinsic motivation) because all users who achieve *Badge (11)* must also achieve *Badge (10)*.

6.4 Results

The empirical results are illustrated in Table 6. We present the independent variables in the first column and the estimated coefficients in the second column. To get an approximation for the base level of effort while users hold a certain badge we calculate the so-called adjusted predictions (AP) or predictive margins for each badge. In our setting, the APs describe how many *Answers* & *Questions* the users would have provided on average if they were to hold a certain badge, while taking into account the impact of all other covariates at the same time. By calculating the differences between the APs we obtain the average marginal effect (AME) for each badge (Cameron and Trivedi 2010). The AMEs reveal how the base level of effort changes after successful badge achievement, which is illustrated in column four. The right side of Figure 1 provides a visual description of the interpretation of the APs and AMEs.

Variables	Model (1)	Model (1) - AP	Model (1) - AME
Badge (2)	0.991** (0.041)	0.885** (0.015)	0.800** (0.033)
Badge (3)	1.725** (0.049)	1.686** (0.035)	1.223** (0.073)
Badge (4)	2.435** (0.058)	2.910** (0.075)	1.154** (0.123)
Badge (5)	2.796** (0.067)	4.064** (0.122)	1.992** (0.218)
Badge (6)	3.132** (0.082)	6.057** (0.215)	2.433** (0.372)
Badge (7)	3.597** (0.094)	8.491** (0.377)	2.481** (0.585)
Badge (8)	3.875** (0.113)	10.97** (0.565)	5.259**** (0.802)
Badge (9)	**4.205**** (0.095)	**16.23**** (0.721)	**5.324**** (1.840)
Badge (10)	**4.931**** (0.232)	**21.55**** (1.884)	**-4.24** (2.732)
Badge (11)	**4.379**** (0.334)	**17.30**** (2.457)	**3.042** (3.078)
Badge (12)	4.261** (0.292)	20.35** (2.470)	1.431 (2.827)
Badge (13)	4.360** (0.181)	21.78** (1.932)	1.321 (2.747)
Badge (14)	4.466** (0.317)	23.10** (2.828)	-1.70 (3.562)
Badge (15)	4.594** (0.229)	21.39** (3.016)	12.16** (3.585)
Badge (16)	5.124** (0.164)	33.56** (2.539)	-2.99 (3.291)
Badge (17)	4.955** (0.219)	30.56** (3.643)	0.541 (4.024)
Badge (18)	4.692** (0.254)	31.10** (3.177)	4.027 (3.928)
Badge (19)	4.886** (0.252)	35.13** (3.521)	-
Relative Progress	1.896** (0.104)	-	-
Relative Progress^2	-0.0135 (0.107)	-	-
Constant	-0.389** (0.021)	-	-
Interaction Terms	✓	-	-
Control Variables	✓	-	-
Individual Fixed Effects	-	-	-
-Ln Likelihood	-4,734,711	-	-
Cluster Robust Standard Errors in Parentheses, * p<0.05, ** p<0.01			

Table 6. Results for the Number of Answers & Questions

We start by analyzing the dynamics in user effort caused by the impact of progress towards a badge. The estimators for the variable *Relative Progress* indicate a positive effect. However, before we can draw conclusions about the impact of progress towards a badge we have to take into account the interaction terms with the badge dummies. Therefore, we calculate an overall AME for the variable *Relative Progress* expressed as semi-elasticity to get a rough approximation for the average effect size. The estimator for the semi-elasticity is 0.015 (0.0003) and significant on a one percent level (z-value=56.04, p<0.01). The estimator indicates that the number of *Answers* & *Questions* increases by 1.5% as a user gets closer to the next badge by 1 (or 1 ppt.). Thus, in line with previous studies on the goal-gradient hypothesis (e.g., Kivetz et al., 2006, Mutter and Kundisch, 2014), we also find a positive effect of progress towards a badge on user effort. Therefore we find support for our first research hypothesis and derive our first result:

RESULT I: Users increase their effort the closer they get to a goal.

After separating the impact of progress towards a badge on user effort, we analyze how the base level of effort is affected by successful badge achievement. In Model (1), the estimators for the badge dummies are positive and significant on a one percent level. Moreover, the size of the estimators increases constantly from *Badge (2)* to *Badge (10).* This increase stops abruptly with the reduction in the level of difficulty for *Badge (11)*. From *Badge (11)* to *Badge (15)* the estimators remain below the estimator for *Badge (10),* and at *Badge (16)* the estimators start to rise again.

For a more nuanced analysis we focus on the APs and the AMEs for the badge dummies which are illustrated in column three and four. All APs are positive and significant on a one percent level. In addition, they reveal the same pattern as the estimators for the badge dummies in the second column. The APs increase constantly from *Badge (2)* to *Badge (10)* and drop suddenly at *Badge (11)* to a level below the AP for *Badge (10).* Over *Badge (13)* to *Badge (15)* the APs fluctuate around the AP for *Badge (10)* and start to increase again at *Badge (16).*

In more detail, the AP for *Badge (2)* is 0.885 which means that users holding *Badge (2)* perform on average 0.885 *Answers* & *Questions* per day. The AP for *Badge (3)* is 1.686 and the difference to *Badge (2)* is 0.800 which means that the users' base level of effort is 0.800 activities per day higher at *Badge (3)* compared to *Badge (2).* This difference is equal to the AME for *Badge (2)*, which is illustrated in column four. The AMEs for *Badge (2)* to *Badge (9)* are all positive and significant on a one percent level. From *Badge (10)* to *Badge (19)* the AMEs turn insignificant except for *Badge (15)*. The emerging pattern highlights that the users increase their base level of effort as long as the level of difficulty of badges increases.

This result is even more explicit when we present the differences between the AP for *Badge (10)* and all other Badges in Table 7. We use *Badge (10)* because this badge represents the most challenging goal on the platform (see Table 2). The AP for *Badge (10)* is significantly larger compared to all preceding badges in the hierarchical badge system including *Badge (9).* The differences in AP for *Badges (11)* to *Badge (15)* fluctuate around zero but are insignificant. The APs for *Badge (16)* to *Badge (19)* are higher than the AP for *Badge (10)*. From our point of view there are two explanations for these high APs from *Badge (16)* to *Badge (19).* First, it might be that these APs are positively affected by self-selection which means that users who hold a more valuable badge are on average more intrinsically motivated to contribute to the community than users who are still in the lower echelons of the hierarchical badge system. Second, it is possible that those users are already aiming for the most valuable badge 'Albert Einstein' (*Badge 20*) and therefore increase their user effort over the remaining four badges.

It is, however, the abrupt slow-down in the rise of the users' base level of effort observed at *Badge (11)* that is most pertinent to our argument. This slow-down occurs simultaneously with the drop in the level of difficulty from 2,500 required status points for *Badge (10)* to only 500 for *Badge (11)*. This indicates that users increase their base level of effort after earning a badge only as long as this achievement is challenging. As predicted by the theory on recurring goals and self-learning, users experience a boost in self-efficacy through the success of goal achievement through the channel of self-learning (users actually learn something about their own capabilities which gives them confidence, i.e. self-efficacy). This translates into a higher base level of effort towards the subsequent goal. The necessary condition is that the success of goal achievement does not come too easily, that is, goals must be increasingly difficult to obtain. The results of our analysis thus mirror almost exactly these predictions of the theory of recurring goals and self-learning, as a user's effort level drops as soon as a goal is not more challenging as compared to the prior goals, in other words, at the transition from badge 10 to 11. Therefore, we find support for our second research hypothesis and we derive our second result:

RESULT II: After successful badge achievement, users increase their base level of effort when attempting to reach the next badge in a hierarchical badge system as long as achieving a badge is challenging.

Differences Estimators	Model (1) - AP	Differences Estimators	Model (1) - AP
Badge (10) - Badge (2)	20.67 $\chi^2(1)$=120, **	*Badge (10) - Badge (12)*	1.20 $\chi^2(1)$=0.2, °
Badge (10) - Badge (3)	19.87 $\chi^2(1)$=111, **	*Badge (10) - Badge (13)*	-0.23 $\chi^2(1)$=0.0, °
Badge (10) - Badge (4)	18.65 $\chi^2(1)$=98, **	*Badge (10) - Badge (14)*	-1.55 $\chi^2(1)$=0.2, °
Badge (10) - Badge (5)	17.49 $\chi^2(1)$=87, **	*Badge (10) - Badge (15)*	0.16 $\chi^2(1)$=0.0, °
Badge (10) - Badge (6)	15.49 $\chi^2(1)$=68, **	*Badge (10) - Badge (16)*	-12.0 $\chi^2(1)$=15, **
Badge (10) - Badge (7)	13.06 $\chi^2(1)$=48, **	*Badge (10) - Badge (17)*	-9.01 $\chi^2(1)$=5.1, *
Badge (10) - Badge (8)	10.58 $\chi^2(1)$=31, **	*Badge (10) - Badge (18)*	-9.55 $\chi^2(1)$=7.0, *
Badge (10) - Badge (9)	5.32 $\chi^2(1)$=8.4, **	*Badge (10) - Badge (19)*	-13.58 $\chi^2(1)$=13, **
Badge (10) - Badge (11)	4.25 $\chi^2(1)$=2.4, °	-	-
Chi-Squared Test, ** p<0.01, * p<0.05, ° p>0.05			

Table 7. Differences Model (1) - AP

6.5 Robustness Checks

We examine a number of robustness checks to demonstrate the robustness of our results. (1) We include the variable *Relative Progress* only in absolute rather than squared terms to try a different model specification; (2) we use only the number of *Answers* per user per day as dependent variable; (3) to account for unobserved time constant heterogeneity we estimate a Poisson fixed-effects model where we include the badge dummies (*(Badge (1)* to *Badge (19)*) and the *Length of Membership* in absolute and squared terms as independent variables; (4) to rule out that our results are driven by outliers we recode the values for the dependent variable number of *Answers* & *Questions* which lie above the 99% quantile with the value of the quantile. Our main results remain qualitatively unchanged for each robustness check.

7 Conclusion

The recent popularity of applying gamification in business applications is based on the premise that gamification offers a range of tools able to motivate user participation in a variety of contexts. With this paper, we enhance the systematic understanding of gamification by investigating how successful goal achievement affects future effort to attain the next goal in a recurring goal framework. We find a positive impact of successful goal achievement on subsequent effort levels to reach the next goal as long as goals are challenging. In line with Drèze and Nunes (2011), we identify self-learning as a key driver for our results. These findings are robust and have survived a range of robustness checks. With these results we contribute to the body of literature on gamification specifically by investigating how goals represented by virtual rewards affect user contribution levels. We also contribute to the literature on recurring goals and self-learning by extending the work of Drèze and Nunes (2011) to virtual reward systems and showing that self-learning in a recurring goal setting also has a positive effect on motivational phenomena such as user effort. As prior research on recurring goals and self-learning has primarily focused on customer retention and loyalty (e.g., Eggert et al., 2015), we are the first to extend this theory to virtual rewards and user effort.

Due to our extension of the theory to virtual rewards, the managerial implications to IS designers are manifold. For example, the popularity and longevity of websites using gamified applications such as StackOverflow or Wikipedia certainly prosper on the demand for content by users, but also on the content contribution by users. In order to motivate users to keep contributing content, the design of the gamification elements that motivate users to contribute content is crucial. Gamification designers should be aware of the positive impact of successful goal achievement in a recurring goal framework and, for the effect to be lasting, to ensure that successive goals are perceived as challenging. Moreover, in terms of structuring a recurring goal system, the prevalence of self-learning would recommend the adoption of multiple goals with an increasing level of difficulty instead of a smaller number of goals that are more difficult to achieve. When setting up a badge system for a website for the first time, it is reasonable to assume that the implementation of our findings into the badge system can be conducted quite easily. For example, designers have to make sure that, ceteris paribus, the necessary status points for a badge are increasing from badge to badge. Our results may be especially, but not exclusively, valuable for long-term goal design. These goals aim at stimulating user effort in the long-run as opposed to short-run concepts like the "starting problem" (Heath et al. 1999), to which our findings are also applicable, nevertheless.

Although our findings are overall consistent with the theory, we recognize that other factors might be at play, which we have not yet accounted for in our research setting, such as the different thematic areas of the platform. While the results from the Q&A community under study may not be directly applicable 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., Krankanhalli et al., 2005, Jeppesen and Frederiksen, 2006). In an environment where individuals are more extrinsically motivated, such as in a Business or Education setting, we would also expect to see a pronounced positive effect of successful badge achievement on efforts to attain the next level. We therefore have reason to believe that this positive effect of self-learning would equally apply to other domains, including Business and Education. Moreover, one might argue that badges in our setting "merely" represent an (arbitrary) translation or aggregation of status points. With regard to this, prior literature (von Rechenberg et al., 2016) provides empirical evidence in support of the claim that goals formed as the aggregate of status points inherit the properties of Prospect Theory's value function and therefore they serve for instance as reference point, attributing user effort into regions of gains and losses (see also Heath et al., 1999).

The impact of successful goal achievement on future effort to attain the next goal in a recurring goal framework represents, in our opinion, a promising avenue for future research. Future work could investigate the relationship between the level of difficulty of a goal and the effect size of goal achievement. Furthermore, future research could analyze the limitations of the recurrence of the positive effect of self-learning over time.

References

Anderson, A., Huttenlocher, D. J. Kleinberg and J. Leskovec (2013). "Steering User Behavior with Badges." In: *Proceedings of the 22nd International Conference on World Wide Web*. Rio de Janeiro: Brasil, 95–106.

Bandura, A. (1982). "Self-Efficacy Mechanism in Human Agency." *American Psychologist* 37 (2), 122–147.

Bandura, A. (1988). "Perceived Self-Efficacy: Exercise of Control Through Self-Belief." In: *Cognitive Perspectives on Emotion and Motivation*. Ed. by J. P. Dauwalder, M. Perrez and N. H. Frida. Dordrecht, the Netherlands: Kluwer Academic, 37–61.

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.

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics using stata*. College Station, TX: Stata Press.

Denny, P. (2013). "The Effect of Virtual Achievements on Student Engagement." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 763–772.

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

Drèze, X. and J. C. Nunes (2011). "Recurring Goals and Learning: The Impact of Successful Reward Attainment on Purchase Behavior." *Journal of Marketing Research* 48 (2), 268–281.

Dzewaltowski, D. A., J. M. Noble and J. M. Shaw (1990). "Physical activity participation: Social cognitive theory versus the theories of reasoned action and planned behavior." *Journal of Sport & Exercise Psychology* 12 (4), 388–405.

Eggert, A., Steinhoff, L., and I. Garnefeld (2015). „Managing the Bright and Dark Sides of Status Endowment in Hierarchical Loyalty Programs." *Journal of Service Research*, 18(2), 210-228.

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 11/26/2015).

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 11/26/2015).

Goes, P., Guo, C., and M. Lin (2016). "Do incentive hierarchies induce user effort? Evidence from an online knowledge exchange" Evidence from an Online Knowledge Exchange" URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2413052 (visited on 05/04/2016).

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.

Heath, C., R. P. Larrick and G. Wu (1999). "Goals as Reference Points." *Cognitive Psychology* 38 (1), 79–109.

Heilizer, F. (1977). "A Review of Theory and Research on the Assumptions of Miller's Response Competition (Conflict) Models: Response Gradients." *The Journal of General Psychology* 97 (1), 17–71.

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

Hull, C. L. (1934). "The Rats' Speed of Locomotion Gradient in the Approach to Food." *Journal of Comparative Psychology* 17 (3), 393–422.

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.

Locke, E. A. and G. P. Latham (1990). *A theory of goal setting & task performance*. Prentice-Hall.

Latham, G. P. and E. A. Locke (1991). "Self-regulation through goal setting." *Organizational behavior and human decision processes* 50 (2), 212–247.

Locke, E. A. and G. P. Latham (2002). "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey." *American Psychologist* 57 (9), 705–717.

Latham, G. P. and E. A. Locke (2007). "New developments in and directions for goal-setting research." *European Psychologist* 12 (4), p. 290.

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

Mutter, T. and D. Kundisch (2014b). Don't take away my status!–Evidence from the restructuring of a virtual reward system. *Computer Networks*, 75, 477-490.

Mutter, T. and D. Kundisch (2015). Behavioral Mechanisms Prompted by Virtual Rewards: The Small-area Hypothesis. In: *Proceedings of the 2015 European Conference on Information Systems (ECIS 2015).* Münster: Germany.

Rauch, M. (2013). "Best Practices for Using Enterprise Gamification to Engage Employees and Customers." In: *Proceeding of the 15th International Conference on Human-Computer Interaction*. Las Vegas, NV, 276–283.

Ryan, T. A. (1970). *Intentional Behavior*. New York: Ronald Press.

Shapiro, C., and H. R. Varian (2013). *Information rules: a strategic guide to the network economy*. Harvard Business Press.

von Rechenberg, T, Gutt, D. and D. Kundisch (2016). "Goals as Reference Points: Empirical Evidence from a Virtual Reward System." *Decision Analysis* Forthcoming.

Wood, R. and A. Bandura, (1989). "Social cognitive theory of organizational management." *Academy of management Review,* 14(3), 361-384.

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