Gamification and engagement with a cycling intervention: the effects of short- and long-term goal setting in promoting behavior change

Anita Lyubenova[1], Fariya Sharmeen[1]

[1]KTH Royal Institute of Technology

Abstract

Gamification, the use of game-design elements in non-game contexts, is one approach to promote cycling as a travel mode. This study explored the effects of short- and long-term goals on cycling behavior in the context of a gamified app with monetary rewards. Weekly challenges (short-term goals) and trophies (long-term goals) were evaluated in a within-subject design across three periods: base gamification, challenges+trophies, and trophies-only. Mixed-effects models showed that adding challenges and trophies to the app increased daily cycled kilometers, but only among men (b=1.4km, p=.003) and regular cyclists (b=1.8, p=.001). Cycling increased further when challenges were removed (1.8km, p=.001), and the effect was not moderated by gender and cycling frequency at baseline. The findings reveal that goals can increase cycling on top of other gamification elements, including monetary rewards. However, the study highlights the limitations of gamified goal setting as a tool to support behavior change among less regular cyclists.

Keywords: gamification, goal setting, cycling, behavior change

1. Introduction

Cycling as a travel mode can contribute both to a sustainable transport system and to public health. However, various barriers limit its wider adoption (Alm & Koglin, 2022), necessitating interventions at different levels to ptomote it. One approach to promote cycling bottom-up, i.e., by increasing individuals' motivation, is to use gamified interventions. Gamification is defined as the use of game design elements in non-game contexts (Deterding et al., 2011). The rationale is to make mundane or tedious tasks more interesting by coupling them with game design elements, known to promote engagement in games. Goal setting is closely related to multiple gamification elements (Krath et al., 2021), however, knowledge about the specific effects of goals within gamification bundles, particularly on cycling behavior, remains limited.

Goals have been considered to regulate human behavior and play a role in behavior change (Locke & Latham, 2002; Michie et al., 2013). Previous research suggests that goal setting in gamification acts in a similar way to goal setting in non-gamified contexts, and even outperform it (Groening & Binnewies, 2019). In gamification, achievements and badges are awarded after reaching a certain milestone, which acts as a goal (Groening & Binnewies, 2019; Gutt et al., 2020). Progress bars, instead, track the advancement towards goals. Furthermore, goal achievement fosters the commitment to subsequent goals with greater difficulty (Gutt et al., 2020), which is a key mechanism underlying level systems of games.

An exisiting gap is, previous investigations of goal setting in gamification did not distinguish between short- and long-term goals. Short- and long-term goals have different characteristics, benefits, and drawbacks and it has been suggested that their combination is more effective than either one of them (Höchli et al., 2018). For example, short-term goals may support the pursuit of long-term objectives by providing immediate, attainable targets and fostering a sense of progress and self-efficacy. On the other hand, the bigger value of a long-term goal can motivate the fulfillment of smaller goals. Such interactions are widely unexplored in gamification setting, even though such knowledge could inform the design of gamified interventions.

The aim of this study was to investigate the individual and combined effects of short-and long-term goal setting on behavior in an app-based gamified setting. The behaviour in this case is cycling for daily activities and the short-term goals were operationalized as weekly challenges, while the long-term goals were signified by milestones (designed as bronze, silver and gold trophies), the progress toward was visualized with a progress bar. To disentangle the individual and combined effects of weekly challenges and trophies a within-subject design was employed, in which participants experienced both elements in one period of time, and only trophies in another period. The study aimed to answer the following research questions:

RQ 1: To what extent did the addition of weekly challenges and trophies to a gamified app affect the daily cycled kilometers?

RQ2: To what extent did removing the weekly challenges affect the daily cycled kilometers compared to the period with both challenges and trophies?

For both questions, exploratory analyses were performed to investigate whether individual differences, such as gender, age, and travel behavior at baseline moderated the effectiveness of the interventions.

2. Methods

The study was conducted as part of the Bike2Green project in Stockholm, Sweden. The project aimed to promote the adoption of daily cycling in Stockholm through a system of reward-based incentives and gamification. To this end, a mobile app, Pin Bike, was employed where participants could log their trips and receive their rewards. Section 2.1 elaborates more on the app and the gamification features.

2.1 Gamification

Participants could log their trips into the app, which were verified by a sensor attached to the bike. Upon completion and validation of a trip, the participants received 2 or 5 SEK per kilometer for non-commuting and commuting trips, respectively. In this way, participants could receive vouchers worth up to 550 SEK (€48) per month that could be redeemed in local shops.

Element	**Target (km)**	**Target (trips)**
Weekly challenges		
Easy	15	4
Medium	25	-
Hard	40	8
Extra hard	60	-
Trophies		
Bronze	0 - 75	-
Silver	75 - 250	-
Gold	250 - 500	-

Table 1. Short- and long- term goals, as represented by weekly challenges and trophies, and their targets, respectively

In addition, points were rewarded per kilometer and were used to determine the position of the participant in the monthly ranking. At the end of the month, the top 40 participants in the ranking received an additional reward of 330 SEK. The app also provided a quantified overview of several targets, including the accumulated monetary rewards, number of trips and cycled kilometers, average kilometers per trip, and saved CO2 emissions.

Weekly challenges set cycling-related goals, such as achieving a specific number of kilometers or completing a certain number of trips within a week and varied in difficulty (Table1). All users were automatically enrolled in the challenges. Upon reaching a target, participants received a congratulatory message and earned points as a reward. Each week there was a mix of easy, moderate or hard challenges, so that each cyclist group had potentially something to strive for.

Trophies designated longer-term objectives intended to be achieved within a 10-week period. These objectives comprised three sequential distance-based milestones, denoted by bronze, silver, and gold trophies (Table 1). The progress bars were updated after each cycling trip, providing real-time feedback on participants' advancement toward the next trophy.

2.2 Study design

In order to disentangle the effects of short and long-term setting from the remaining gamification features, we conducted a field experiment with within-subject design. The study period between August 19 and November 29, 2024 was split into three parts, with different gamification characteristics active in each part, as specified in Table 2. Period 1 (P1, 6 weeks) was the control period, not including any goal setting. Period 2 (P2, 6 weeks) introduced both short- and long-term goals through weekly challenges and progress bars toward trophies. Period 3 (P3, 3 weeks) removed the weekly challenges while retaining the possibility to attain trophies.

Gamification elements	**Period 1** (w. 34-39)	**Period 2** (w. 40-45)	**Period 3** (w. 46-48)
Points	+	+	+
Kilometric rewards in vouchers	+	+	+
Monthly ranking	+	+	+
Monthly reward for top users	+	+	+
Performance stats (trips, km, etc.)	+	+	+
Weekly challenges		+	
Progress bars toward trophies		+	+

Table 2. Periods (and duration in week numbers) and the availability of the gamification elements

Data

The inclusion criteria ensured participants had sufficient data during the first period and were active app users at the start of the second. The dataset included all recorded trips (date and distance) per participant, daily weather data (temperature, precipitation, wind speed) from SMHI, and baseline survey data on socio-demographics and mobility behavior. Mobility behavior, based on mode frequency questions (car, bike, walk, public transport) rated on a 5-point Likert scale, was imputed for missing data using chained random forests and reduced to three mobility classes via latent class analysis: people who use exclusively bike ("bike only"), those who use almost only public transport ("public transport"), and those who use bike often (3-4 times per week) but also walk and use public transport from time to time ("bike mixed").

Statistical analysis

The study examined the effect of gamification elements on daily kilometers cycled, using mixed-effects linear regression models to account for the hierarchical data structure, with daily kilometers nested within participants. The analysis controlled for within-person (e.g., temperature, precipitation, windspeed, periods) and between-person (age, gender, car ownership, travel behavior) variability. Models were built hierarchically, starting with a random intercept-only model to compute the intraclass correlation coefficient (ICC) and progressing to random slope models with interaction effect to assess variation in intervention effects across participants. The reported coefficients were unstandardized (b) alongside their corresponding p-values, with a significance threshold of 0.05. R^2 was computed as an indicator of the explained variance of the model for both the lower (days) and cluster (individuals) level, by using the intercept-only model as reference. Model fit was evaluated using Akaike Information Criterion (AIC), Bayesian Information Criterion

(BIC), and likelihood ratio tests. All analyses were performed in R (v4.3.3) using the packages lme4 (v1.1.35.5) and lmerTest(v3.1.3)

3. Results

3.1 Descriptive statistics

Of the 1047 individuals who registered in the app between March and November 2024, 866 recorded a trip since August 19 and had complete app user data (Table 4). There was a slightly higher proportion of males (55%) compared to females (43%). Most participants had advanced education, with more than 60% holding a master's or a doctorate degree. The majority of the sample (62%) had access to a car, even though car frequencies were very low in the sample. Cycling was the most frequent mode of transport, and a small part of the sample used public transport as their main mode (15%).

The recorded cycling took somewhat different patterns over time across the three mobility classes

Variable	**N (%)**	**Daily km (M,SD)**
Gender		
Male	480 (0.55)	17.3 (12.8)
Female	376 (0.43)	14.6 (9.9)
Other	10 (0.01)	12.5 (7.5)
Education		
Primary school	14 (0.02)	15.6 (14.4)
Secondary school	122 (0.14)	19.3 (15.4)
Bachelor's degree	210 (0.24)	16.3 (11.9)
Master's degree	403 (0.47)	15.0 (10.1)
PhD	117 (0.14)	16.4 (11.6)
Car access		
Yes	540 (0.62)	17.3 (12.8)
No	326 (0.38)	14.5 (9.9)
Mobility class		
Bike only	510 (0.59)	16.2 (12.2)
Bike mixed	223 (0.26)	16.4 (10.9)
Public transport	133 (0.15)	15.9 (11.0)
Total	866 (100%)	16.2 (11.8)

Table 4. The sample's socio-demographic characteristics and recorded cycling behavior. Abbreviations: M = mean, SD = standard deviation, N = number of participants

(Figure 2). The "bike only" class shows a pronounced increase in daily kilometers for some individuals during P2 compared to P1, suggesting a potential impact of the additional gamification elements. In contrast, "Bike mixed" and "Public transport" appear to maintain more consistent patterns across the periods, with less visible response to the gamification interventions. Notably, the "Bike mixed" class was characterized by clear weekday-weekend distinction, with regular trips during weekdays and fewer trips during weekend. This difference was less pronounced in the "bike only" and the "public transport" classes, where trips were more sporadic.

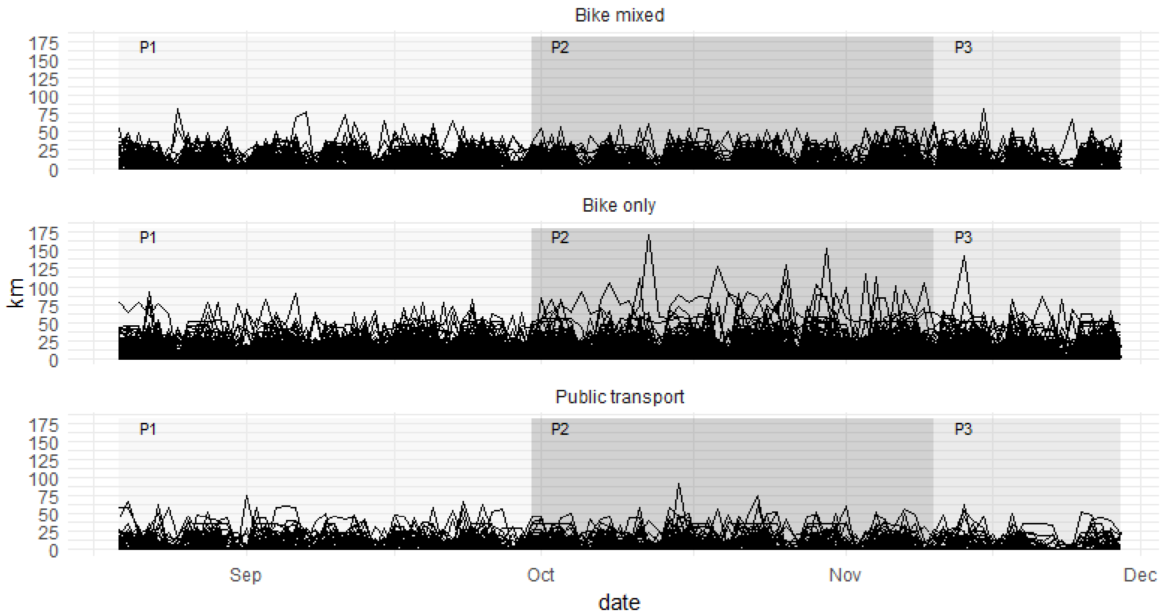

Figure 2. Recorded kilometers for each day, grouped by mobility class, where each line represents an individual. The shaded areas indicate the study periods: period 1 (P1) base gamification, period 2 (P2) with trophies and weekly challenges, and period 3 (P3) with trophies only.

3.2 Trophies and challenges

During P2, the addition of weekly challenges and progress bars significantly increased daily cycled kilometers, with participants cycling an average of 1.60 km more per day compared to P1 ($p<0.001p<0.001$), while controlling for weather, age, gender, and mobility class (Table 5, Model 1). Car access was not a significant predictor ($b=-0.1$, $p=0.84$) and was excluded from subsequent models for parsimony. The intraclass correlation coefficient (ICC) was 0.21, indicating that one-fifth of the total variation in cycling behavior was attributable to individual differences, with the remaining variance resulting from day-to-day fluctuations.

Model 1 with a random slope for P2 significantly improved model fit of the preceding random intercept model ($\Delta AIC=888$, $\Delta BIC=871$, $p<0.001$), suggesting heterogeneous effects of the gamification intervention across participants. In Model 2, interactions between P2 and individual characteristics revealed that only mobility class significantly moderated the effect of P2, such that "bike only" cycled 1.62 km more during P3 than "bike mixed" ($p=0.001$), though this moderation explained just 4% of the variance in slopes. Notably, the main effect of P2 became insignificant ($b=0.47$, $p=0.34$), indicating that the intervention's effectiveness was primarily driven by participants

	Model 1		Model 5	
	Coef.	**p-value**	**Coef.**	**p-value**
Fixed effects				
Intercept	0.67	0.250	0.94	0.112
P2	**1.60**	**<0.001**	0.47	0.349
Temperature	**0.25**	**<0.001**	**0.25**	**<0.001**
Precipitation	**-0.02**	**0.029**	**-0.02**	**0.027**
Windspeed	**-0.29**	**<0.001**	**-0.29**	**<0.001**
Age	**0.13**	**<0.001**	**0.13**	**<0.001**
Gender				
Male	**1.43**	**0.002**	**1.37**	**0.003**
Other	-0.48	0.815	-0.14	0.948
Mobility class				
Bike only	**2.19***	**<0.001**	**1.80**	**0.001**
Public transport	1.24	0.080	1.13	0.119
P2 x age			0.005	0.783
P2 x gender				
x Male			0.21	0.61
x Other			-1.43	0.46
P2 x Mobility class				
x Bike only			**1.62***	**0.001**
x Public transport			0.45	0.499
Random effects				
σ_e^2	82.90		82.90	
σ_{00}^2	19.17		19.14	
σ_{11}^2	14.19		13.64	
Model characteristics				
R_{L1}^2	0.05		0.05	
R_{L2}^2	0.19		0.19	
AIC	267119		267115	
BIC	267238		267277	

Table 5. Model 1 and 2 are random-intercept random-slope models regressing the daily kilometers on P3 and control variables. Reference categories were female, and bike mixed mobility class. $N_{participants}$ = 435, $N_{days\ x\ individuals}$= 26047; σ_e^2 – residual variance at the lowest level 1 (days), σ_{00}^2 – variance at cluster level (individuals); σ_{11}^2 – slope variance; R_{L1}^2 – explained variance at the lowest level; R_{L2}^2 – explained variance at cluster level; coefficients in bold were significant at the 0.05 level.

in the “bike only” class. However, BIC is larger in model 2 (ΔBIC = -162), meaning that the improvements in model fit are not enough to justify the added complexity. Overall, the models explained little of the day-to-day variance (5%), and nearly one-fifth of the inter-individual variance.

Variable	Model 3 coef.	p-value	Model 4 coef.	p-value
Fixed effects				
Intercept	**4.23**	**<0.001**	**3.68**	**<0.001**
P3	**1.05**	**<0.001**	**1.77**	**0.001**
Temperature	**0.2**	**<0.001**	**0.2**	**<0.001**
Precipitation	**-0.06**	**<0.001**	**-0.06**	**<0.001**
Windspeed	**-0.46**	**<0.001**	**-0.46**	**<0.001**
Age	**0.16**	**<0.001**	**0.13**	**<0.001**
Gender				
Male	**1.48**	**0.004**	**1.79**	**0.004**
Other	-1.68	0.475	-0.69	0.805
Mobility class				
Bike only	**2.65**	**<0.001**	**3.03**	**<0.001**
Public transport	0.87	0.365	1.85	0.104
P2 x age			0.03	0.122
P2 x gender				
x Male			-0.40	0.348
x Other			-1.29	0.508
P2 x Mobility class				
x Bike only			-0.49	0.341
x Public transport			-1.27	0.109
Random effects				
σ_e^2	98.89		98.89	
σ_{00}^2	35.29		35.12	
σ_{11}^2	10.98		10.68	
ICC	0.24		0.24	
Model characteristics				
R_{L1}^2	0.03		0.03	
R_{L2}^2	0.12		0.13	
AIC	195182		195185	
BIC	195297		195340	

Table 6. Model 4 and 5 regressing the daily kilometers on P3 and control variables. Reference categories were female, and bike mixed mobility class. $N_{participants}$= 435, $N_{observations}$= 26047; σ_e^2 – residual variance at the lowest level 1 (days), σ_{00}^2 – variance at cluster level (individuals); σ_{11}^2 – slope variance; R_{L1}^2 – explained variance at the lowest level (days) ; R_{L2}^2 – explained variance at cluster level (individuals); coefficients in bold were significant at the 0.05 level.

3.3 Throphies only

During P3, even though weekly challenges were removed, participants cycled an average of 1.05 km more per day compared to P2 (p<0.001), while controlling for weather, age, gender, and mobility class (Table 6, Model 3). Similar to model 1, car access was not significant. The IIC was 0.24, with almost a quarter of the variation being at individual (cluster level).

The random slopes included in model 3 significantly improved the fit compared to the random-intercept model (ΔAIC=365, ΔBIC=344, p <.001), indicating heterogeneity in the individuals' responses to P3. However, none of the investigated variables in model 4 – age, gender, and mobility class – significantly moderated the effect of P3 - and thus - explained that heterogeneity. In result, model 4 was also not superior to model 3 according to BIC and significance testing (p=0.12). Both model 3 and 4 explained little variance at the day (3%) and individual level (12% and 13%).

4. Discussion

This study investigated the effects of short- and long-term goals on cycling behavior in a gamified setting, operationalized as weekly challenges and trophies, respectively. The combination of challenges and trophies led to a significant increase in daily kilometers cycled, although this effect was observed only among men and the most regular cyclists. This highlights a paradox – we aimed to promote cycling but the only ones who benefitted from the intervention were the individuals who were already frequent cyclists.

These findings align with previous research which demonstrated that cycling increased among individuals who opted to participate in challenges (Huang et al., 2021). However, the majority of participants in that study did not choose to engage with the challenges, and those who did were subject to self-selection. Our study extends this understanding by revealing that even without requiring participants to "sign up" for specific challenges, only a subset of individuals responded to the goal-setting intervention. Although we were unable to characterize this subgroup in detail, it is plausible that performance-oriented cyclists found the goal-setting features more appealing compared to commuters, whose cycling behavior is likely influenced more by daily activities and practical necessities than by external motivational goals.

Interestingly, the removal of challenges (P3) was associated with an increase in cycled kilometers compared to P2. This outcome may be attributed to the perception that trophy targets became more attainable, motivating participants to exert additional effort to achieve them. This interpretation is consistent with the goal-gradient hypothesis, which suggests that motivation and effort intensify with proximity to the goal (Mutter & Kundisch, 2014). Thus, the value of short-term goals may be reduced the closer and more attainable a long-term goal is.

A limitation of the study is that participants were required to manually start and end the tracking of each trip, a process that may have been perceived as tedious and could have led some participants to forget to log their trips. This raises the possibility of an alternative explanation: rather than increasing their cycling activity, participants may have been more motivated to log their trips consistently. While the presence of monetary rewards likely provided sufficient motivation to mitigate this issue, such an explanation cannot be entirely ruled out based on the available data. Another limitation is that the majority of our sample consisted of people who already cycle regularly. Therfore we are cautious to draw definitive conclusions about the effectiveness of this gamified intervention in the general population of people who are not already motivated to cycle.

In conclusion, this study demonstrates that goal setting can effectively encourage increased (recorded) cycling. However, caution is warranted to account for self-selection sample biases and when applying this approach to promote cycling among non-regular or non-competitive cyclists. For these individuals, the perceived value of achieving performance-based goals may be lower, suggesting that goal setting may not serve as a primary motivating factor for increasing their cycling activity.

References

Alm, J., & Koglin, T. (2022). (In)capacity to implement measures for increased cycling? Experiences and perspectives from cycling planners in Sweden. *Journal of Urban Mobility*, *2*, 100029. https://doi.org/10.1016/j.urbmob.2022.100029

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. ACM. https://doi.org/10.1145/2181037.2181040

Groening, C., & Binnewies, C. (2019). “Achievement unlocked!” - The impact of digital achievements as a gamification element on motivation and performance. *Computers in Human Behavior*, *97*, 151–166. https://doi.org/10.1016/j.chb.2019.02.026

Gutt, D., Rechenberg, T. von, & Kundisch, D. (2020). Goal achievement, subsequent user effort and the moderating role of goal difficulty. *Journal of Business Research*, *106*, 277–287. https://doi.org/10.1016/j.jbusres.2018.06.019

Höchli, B., Brügger, A., & Messner, C. (2018). How Focusing on Superordinate Goals Motivates Broad, Long-Term Goal Pursuit: A Theoretical Perspective. *Frontiers in Psychology*, *9*, 1879. https://doi.org/10.3389/fpsyg.2018.01879

Huang, B., Thomas, T., Groenewolt, B., Claasen, Y., & van Berkum, E. (2021). Effectiveness of incentives offered by mobile phone app to encourage cycling: A long-term study. *IET Intelligent Transport Systems*, *15*(3), 406–422. https://doi.org/10.1049/itr2.12034

Krath, J., Schürmann, L., & Korflesch, H. F. von (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, *125*, 106963. https://doi.org/10.1016/j.chb.2021.106963

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037//0003-066x.57.9.705

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine : A Publication of the Society of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Mutter, T., & Kundisch, D. (2014). Behavioral mechanisms prompted by badges: The goal-gradient hypothesis. https://core.ac.uk/download/pdf/301363344.pdf