

Effectiveness of incentives offered by mobile phone app to encourage cycling: A long-term study

Bingyuan Huang¹ | Tom Thomas¹ | Benjamin Groenewolt² | Yorick Claasen¹ | Eric van Berkum¹

¹ Centre for Transport Studies, University of Twente, AE Enschede, the Netherlands

² Gemeente Enschede, AE Enschede, the Netherlands

Correspondence

Bingyuan Huang, Centre for Transport Studies, University of Twente, P.O. Box 217, 7500 AE Enschede, the Netherlands.
Email: b.huang@utwente.nl

Abstract

Reduction of car use is one of the most effective ways to tackle congestion-related problems. Using positive incentives to stimulate bicycle use is one possibility to reduce car use. Cycling is a sustainable transport mode that uses little space and is healthy. There is evidence that positive incentives may be more effective than punishing travellers for undesirable behaviour, and the emergence of mobile applications for delivering interventions has opened up new opportunities for influencing travellers. So far, few studies have focused on exploring the effectiveness of positive incentives on long-term behavioural change. We used the SMART app to deliver positive incentives to more than 6000 travellers in the Dutch region of Twente. The app automatically tracks users and provides incentives such as challenges with rewards, feedback, and messages. This study covers the period from March 2017 to June 2018, in which more than 1000 SMART users participated in monthly challenges. We evaluated the effects of the challenges and rewards and found that the challenges did encourage cycling and reduced car use in the short term. There is also some evidence for behavioural change over a longer time period.

1 | INTRODUCTION AND BACKGROUND

Traffic and its externalities, such as the emission of greenhouse gases, is increasingly causing problems for almost all major cities. One of the challenges for transportation researchers is to change people's travel choices and getting them to use sustainable transport modes. The negative health effects of a sedentary lifestyle provide another compelling reason to encourage active transport modes. In the Netherlands, people cycle about 10% of all trip kilometres, which is a significantly larger fraction than in other countries. The Dutch government wants to encourage cycling even more and aims to increase the cycled distance by 20% between now and 2027 [1].

Offering positive incentives is a relatively cheap way to encourage cycling because it does not require infrastructure investment. Voluntary travel behavioural change (VTBC) schemes that use incentives such as rewards, feedback, subsidies and public transport discounts can result in a shift from

car use towards more sustainable travel modes, for example, the work of Brög et al. [2], Sanjust et al. [3], Ben-Elia and Ettema [4], and Lachapelle [5]. Such schemes work better than fiscal measures in the sense that they do not encourage socioeconomic inequity [6].

Traditional VTBC solutions require person-based interaction either by phone or home interviews, which is inherently expensive and may induce biases stemming from social interaction and communication. The emergence of the social web and mobile applications for delivering interventions offers opportunities to reduce the costs and enables the use of ICT-based persuasion technology for influencing travellers [7]. Table 1 lists eight mobile app projects intended to encourage VTBCs [8, 9, 18–20, 10–17]. The app design of those studies followed or fitted in the persuasive systems design model [21–24], which offers a way to analyse, design and evaluate the persuasion context. Common persuasive design features used in those studies were personalised feedback, self-monitoring, challenges and goal setting, social comparison, and rewards. Field experiments

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Intelligent Transport Systems* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

TABLE 1 Overview of software and information systems based interventions for travel behavioural change

Study	Sample sizes	Experiment period	Result	Credibility system
UbiGreen 2009 [8]	14	Four weeks	As an early-stage study, it cannot reasonably be expected to confirm that behaviour change occurred. The qualitative results indicate that participants did start new behaviours	Not functioning optimally
Tripzoom, SUNSET project 2012 [9, 10, 11]	108 Enschede, NL 6 Leeds, UK 95 Gothenburg, SE	Multiple experiments: Each lasted from two weeks to two months.	The main findings support the effectiveness of the incentives to encourage travel behaviour change.	The clear system architecture of the app, security layer, third-party services, city authorities, High accuracy of trip, modes, location recognition (need improvement), around 25% total corrections
SUPERHUB app 2013 [14]	8	Four weeks	A modest increase (14%) in 'sustainable transport choices'	No automatic tracking data, No detail of support systems
Quantified traveller (QT) 2013 [15, 16]	135	Three weeks	QT is useful to significantly reduce the car mileage and to a lesser extent to encourage walking/cycling	Clear system architecture, third-party services, High accuracy of trip, modes, location recognition (need improvement). 13.5% total corrections
MatkaHupi app [13] SUPERHUB project 2013	8	Four weeks	Challenges were, in general, perceived positively	No detail of support systems
Peacock project [12, 18] 2015	37	Eight weeks	Findings show that persuasive approaches have the potential to influence users' attitude towards sustainable mobility	Clear system architecture, More than 40% total corrections, indication low detection accuracy, Bugs and app crash problems
Metropia app Los Angeles pilot 2015[17]	36	10 weeks	20% travel time can be saved for people who changed their travel behaviour	High accuracy. Government authorities
OPTIMUM project [20]	76	Six weeks	Apps were used to reward sustainable transport, including cycling. The amount of cycling respectively almost doubled or even more than doubled as a result	No detail of support systems

provide evidence that these features are important for influencing users to change their travel behaviour; however, clear results on their effectiveness based on field evaluations are still missing (Table 1). Most evaluations of behaviour change interventions proposed in this area are rather short term, involve small groups of participants, and provide limited evidence of lasting behavioural impact [24, 25].

This study focuses on exploring positive incentives on behaviour change, especially for modal shift from car to bike, by using a mobile app (SMART app) in Enschede, the Netherlands. The new travel app includes persuasive design features such as feedback, travel information, self-monitoring, nudging and gamification elements. Due to the collaborative feature of VTBC-based travel apps, it is critical to develop the travel app or the experiment to get broad engagement and run for a long time. A better grasp of the behaviour change for the long term by using the mobile app will aid authorities and private entrepreneurs to design effective and appealing apps, eventually translating into a broader potential of VTBC.

Mobility is a highly habitual activity, which is not easily interrupted, except for short periods [26]. Many studies also high-

light that long-term involvement more often leads to positive results [26–28]. Several studies are optimistic about gamification and its potential to extend the commitment of the user [11, 27, 29–32]. That gamification would have a positive impact on participants' involvement is also supported by Seaborn and Fels [33]; however, they also implied that there is a lack of longitudinal study designs. Thus, more research is needed to discover the gamification effect on app involvement and commitment and its impact on behaviour change, especially under a long-term study period.

Different gamification strategies were involved in current travel apps. For rewards, tangible rewards such as money [20] [34] and in-kind rewards, i.e. points [17, 9, 10, 11, 32], and non-tangible rewards, i.e. praise [14, 12, 18] were often used in current apps. Based on our previous study, in-kind rewards are suggested for travel apps. In reality, by having a web shop associated with in-kind rewards may make people more enthusiastic about the in-kind reward than about the money, because money has an impersonal character, which would decrease the feeling of enjoyment. [35–37]. Challenges and goal-setting are also essential features in gamification. Many goal-setting studies showed

that specific and challenging tasks led to higher performance than easy or “do your best” tasks [38–40], however, current apps for cycling promotion reward users by directly calculating the kilometres they cycled, without a specific challenge, or massive challenges that are not suitable for every user. On the other hand, setting specific goals for long-term projects may be disadvantageous for the user participation, as users immediately tend to dropout upon finishing that goal [41]. Moreover, before considering personalised challenges or challenges that explicitly target a certain social-economic segment of travellers, it should be acknowledged that in many cases targeting specific segments or personalised targeting may be costly and even ethically troublesome. Hence, understanding of how to deal with challenges of involving users is important.

App credibility is an essential persuasive element which was often not achieved in mobile phone related studies. Ignorance or technical flaw may directly impact recruitment and engagement rate. Gamification may increase involvement and support long-term interest, but app credibility can be more crucial to lead a successful app. The current travel mobile apps are often related to technology, privacy and reliability issues (Table 1, column ‘credibility system’) [42], which can decrease the credibility of a mobile app [43–46]. Since users will engage with apps that they perceive as credible but navigate away from those they do not consider credible, high detection accuracy and also security features are therefore important so that the app does not award incentives unfairly or inappropriately. The greater the app’s accuracy and security, the more users it will attract. Besides, social trust is important as a motivator for sustainable behaviour change since trust reinforces peoples’ engaging behaviour, that is, acceptability and public involvement [47, 48]. Therefore, third-party endorsements, authority and surface credibility to increase the trust between individuals and the community at large can also enhance app credibility [43]. Thus, a travel app engaging in encouraging sustainable behaviour for society need multidisciplinary cooperation, such as local authorities to manage, commercial parties to support rewards and IT company to support technology.

Given the gap in current literature, this study contributes to the body of knowledge by designing a monthly self-chosen challenge and by analysing travel behaviour tracked by the SMART app in a real-world environment for over a year, supported by local government, commercial parties and IT companies. The main aim of this study is to analyse and explore the app users’ changes in travel behaviour and changes in interest and commitment, both in a short and long term; as well as to identify key factors that influence the response to gamification (challenge and reward), and to explore whether the type of challenge played a role in the change in behaviour. The hypotheses related to the objective are proposed as follow:

1. H1: The app has broad recruitment and high retention, and the user who has high challenge commitment stays longer in the app.
2. H2: The challenges and rewards as gamification mean motivate travel behaviour change for a short term and travel

patterns; users challenge commitment impact behavioural change.

3. H3: There is long term sustained behaviour change but for users with high engagement.

The study is organised as follows: Section 2 provides methods about the SMART app, the case study and introduces the data. Section 3 presents the results, and finally, Section 4 presents the discussions, limitations and concludes the study.

2 | METHODS

This research is based on the SMART application, which records the users’ mobility on their mobile phones. Everyone can download and use this app and choose to participate in monthly cycling-related challenges.

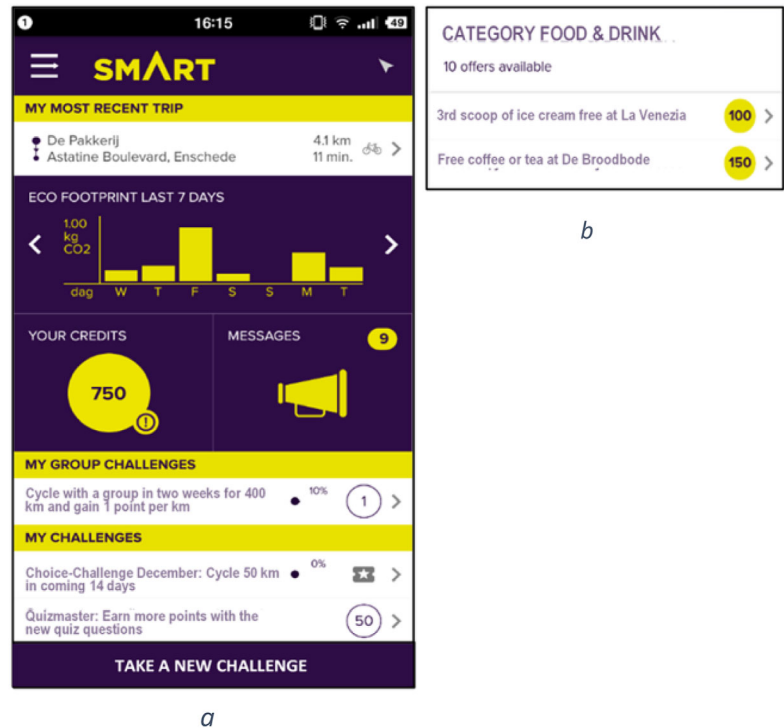
2.1 | SMART app

The SMART app attempts to nudge travellers towards sustainable transport modes, especially cycling, by providing incentives [49]. It runs on the Android and IOS platforms and records the users’ trips. We designed our interventions by drawing on other mobile app-based studies, which aimed to inspire VTBC and contained common persuasive design features (personalised feedback, self-monitoring, challenges and goal setting, social comparison, rewards and praise, traffic information and travel suggestions) [50, 51]. The cooperation between different stakeholders (among which the ICT company behind the app, the city of Enschede, the University) supports the app’s credibility.

Figure 1(a) shows the SMART app’s dashboard, from which users can explore all functions. The following persuasive strategies are used.

1. Self-monitoring and feedback: On the SMART home page and ‘my mobility’ page, users can see their trip history per travel mode and daily CO₂ emission.
2. Challenge and goal setting: In SMART, the living lab operator (municipality) sets up challenges to promote cycling. Users who download the SMART app need to select and join challenges on the challenge page (‘take a new challenge’ in Figure 1(a)), which shows all available challenges. Users are able to choose any suitable challenge. When a user accepts a challenge, the system starts to keep track of the targeted behaviour and shows the progress. This is done by continuous tracking (using GPS, accelerometer data etc.) and by using advanced algorithms that combine travel speeds and routes to determine the transportation mode the traveller is using.
3. Rewards and praise: Rewards are provided upon completion of the challenge. The SMART app starts tracking the user’s target behaviour when a challenge is chosen. When the challenge is completed, the system immediately awards the corresponding number of points, which is also shown in the dashboard. On the ‘reward’ page, the earned points

FIGURE 1 SMART app screenshots. (a) Left: Dashboard, (b) right: System credibility and third-party web shop



can then be redeemed towards discounted products and services (Figure 1(b)). Multiple stakeholders joined the project to support the web shop in the app, which created a win-win, for the municipality (funding) and for the participating shops (sales). Huang et al. [37] concluded that in-kind gifts from a web shop have a more positive impact than cash rewards and may ultimately make users more satisfied [35]. Additionally, the SMART app is able to pop up notifications to praise the users if they travel sustainably.

4. Traffic information and suggestions: Event and traffic information are also offered through messages ('messages' in Figure 1(a)). SMART is able to give useful information about the actual local traffic situation and notifies users of road works or large-scale events that lead to extra traffic. Based on this, SMART may also suggest travel alternatives to enable the users to optimise their travel plans.
5. System credibility: System credibility provides users with credible and authentic information so that it decreases the dropout rate. The SMART app is updated regularly and maintained by Mobidot and supported by the city of Enschede. Moreover, its users are informed in advance about the app's privacy protection, and the travel mobility data is updated every day. Finally, new challenges are updated regularly and the shops where vouchers can be redeemed are all well known (Figure 1(b)).

2.2 | Recruitment

Enschede (the municipality in question) used Facebook, flyer actions on the streets, digital advertisements, and cooperation with local employers to promote the SMART app, aiming to

help the traveller to minimise travel time and cost or/and travel sustainably and healthily. To decrease the risk of selection bias, we emphasised that the SMART app can help travellers to travel smarter and make them more aware of the region.

The recruited users were not asked to provide extra information (such as age or gender) and could participate in challenges immediately. In other words, they were not recruited for an experiment but they can simply use the app if they wanted to. This way, we created a realistic real-life context to analyse the travellers' true behaviours. However, as the users could immediately use all functionalities of the SMART app, we were not able to do a 'before' measurement, which is an important drawback of this study.

2.3 | Case study: Monthly choice challenges

Based on literature reviews [38–41], we aimed to design a challenge that is specific to finish in a time period and is not too easy but also not too specific in a long term to prevent users from losing interest and thus drop out. Moreover, considering cosy, we wanted to avoid personalised challenges. Based on the above considerations, we designed a self-chosen challenge that comes every month, which is named 'monthly choice challenge'.

From March 2017, SMART users could join the 'monthly choice challenge'; those who joined were challenged to accomplish a certain cycling distance or frequency or cycle to a certain location. The monthly choice challenge is still ongoing, but we were allowed to use data from March 2017 until June 2018 for this study. Every month, one monthly choice challenge is offered. In the first two weeks of each month, the SMART users can choose one challenge out of five or six; the challenge

TABLE 2 Monthly choice challenge options and rewards

Challenge types	Challenge options (related reward in points)	Month
Frequency challenge	The number of bike trips to be made within 14 days	May, August 2017; June 2018
	Choice: 2, 5, 10, 25, 50 trips	
	Rewards 60, 90, 150, 330, 630.	
Distance challenge	The number of bike kilometres to be ridden within 14 days	December 2017; May 2018
	Choice: 5, 20, 50, 100, 200 km	
	Rewards 50, 110, 230, 430, 830.	
	The number of bike kilometres to be ridden within xx days	January 2018
	Choice: 30, 60, 90 km in 14 days or 60 km in 3 or 7 days	
	Rewards 150, 270, 390, 510, 390	
Mixed challenge: Frequency and distance	The number of bike trips or kilometres to be made within 14 days	April, July, October 2017; April 2018
	Choice: 5, 10, 50, trips, or 5, 50, 200 km	
	Rewards 90, 150, 630, or 50, 230, 830	
Minimum distance challenge	The number of days to ride a bike for at least 10 kilometres	February 2018
	Choice: 1, 3, 6, 9, 12 days	
	Rewards 70, 150, 270, 390, 510	
	The number of bike kilometres to be ridden on at least 10 out of 14 days.	November 2017 March 2018
	Choice: 1, 2, 5, 10, 15 km per day	
	Rewards 70, 110, 230, 430, 630	
Location	The number of bike trips to go to a certain location within 14 days	March, June, September 2017

immediately starts when a SMART user clicks to join the challenge. The duration of the challenge is up to two weeks; participants can complete their challenge in less time.

We introduced different types of choice challenges: Frequency challenges, distance challenges, mixed challenges (combined frequency and distance), minimum-distance challenges (on certain days, cycle a minimum distance) and location challenges. The five different types of challenges were designed and distributed throughout the year to see which type is more effective. Rotating challenges may also encourage users to keep participating [52].

Table 2 lists the options and related rewards. The more difficult the challenge that the user chooses, the more points the user can gain by completing that challenge. The relation between challenge difficulty and reward is a simple linear function with a constant (baseline points) to ensure that users who rarely cycle can still collect a reasonable number of points if they cycle a short distance or a few times. Choosing a difficult challenge means more rewards; however, if the participant fails to complete the challenge within the challenge period (two weeks), he or she will not receive any points. Participants therefore have to choose between an easy challenge with a relatively small reward but a greater chance to be successful or a difficult challenge with a relatively high reward but a greater chance of failing. Participants immediately are awarded if they meet their chosen challenge in time. They can redeem the points in the app's web shop (see Figure 1(b)).

A post-survey was sent through the SMART experience sampling question service to collect more information on the rea-

sons why users did or did not change their behaviours. The post-survey questionnaire was sent out immediately after completion of the challenge or after the challenge period if participants did not complete the challenge.

2.4 | Data

2.4.1 | Data resource

We used SMART data between March 2017 and June 2018. The data consists of trip data, post-survey data, and information about when users started and completed a particular challenge. The trip data contains the origin, destination, departure time, arrival time, transport mode, activity, average speed, and trip distance. The transport mode and trip activity are determined in the back-end by processing and learning algorithms (see [53]). When a user uses the functions of the SMART app, trigger data are generated. These contain the user's ID, type of trigger (challenge, survey question, message), the status of trigger (accepted, answered, awarded, read), response to the trigger (number of points awarded, reply to a survey question), trigger date and trigger time.

2.4.2 | Indicators and segments

The indicators we chose to evaluate behavioural change are distance-based mode share and daily distance per mode. We

inspected these indicators by aggregating data on a monthly level. We chose a period of a month because a new set of choice challenges was offered each month. We thus had data points per month for each user ('person-month data'). The distance-based mode share is the percentage of the total distance that is covered by each mode during the measurement period. We used distance rather than trip frequency because the latter might be disproportionately influenced by people shifting from walking to cycling for the shortest trips. The daily mode distance is the total distance covered by a certain mode divided by the number of days in the measurement period. We considered several measurement periods per month, that is, the whole month, the period during the challenge, the period outside the challenge period. Note that the latter two are variables among users, dependent on when participants started and completed a challenge. In the long-term analysis, we also used the last week of each month (corresponding with days outside the challenge period for almost all users) to determine the sustained effect of the campaign.

We also performed segmentations to analyse differences between challenge and trip types. As mentioned earlier, we did not have data on personal characteristics of users. However, we made a distinction between users who chose easy challenges and users who chose difficult challenges and checked whether behavioural changes occurred in both groups. We defined the difficulty of the challenge by comparing the challenge against the user's behaviour in the previous month. In other words, if a challenge required the user to cycle more than in the previous month, we assessed the challenge as (relatively) difficult. By contrast, if a challenge required less effort than before, we considered the challenge (relatively) easy. In addition, we distinguished between different types of trips. For example, commuting behaviour may be more difficult to change than recreational travel behaviour [54]. More generally, we assumed that habitual or regular trips may be more difficult to change. If more than eight trips were made between the same origin and destination in one month, then we considered these trips regular trips. This criterion was chosen to obtain more or less equal sample sizes per group.

2.4.3 | Data selection

In total, 6214 users downloaded and used the SMART app between March 2017 and June 2018, with 23,233 person-months of data. To make sure that the tracking data from the SMART app reflected the real travel behaviour, we excluded person-months that had fewer than 15 days of recorded trips. Moreover, only trips below 20 km were counted for travel behaviour analysis, as cycling is not a viable option beyond 20 km [32]. These trips cover almost all trips within the urban area of Enschede (and the neighbouring city of Hengelo) and therefore can be regarded as urban trips. By selecting these trips, we also neutralised the disadvantage of distance-based mode share in which non-recurrent long-distance trips have a disproportionately large effect on mode shares.

After error-checking, cleaning and filtering the data based on the above criteria, 5525 users remained with data from the total

22,174 person-months. In total, 1868 out of 5525 users joined the monthly choice challenge at least once. We call those users 'participants'; they make up the 'experiment group'. The other 3657 users never joined a challenge but used the SMART app regularly for other incentives, such as trip history, and travel information. They are in the No Challenge group. The 1868 participants had 11,076 person-months in total, but the participants only participated in the monthly choice challenge about every other month on average (50%). To check whether participants perhaps only used SMART during challenges and turned off the app to extend battery life after a challenge, we compared the days with recorded trips inside and outside challenge periods. Figure 2(a) shows the average proportion of days with trips recorded. Figure 2(b) displays the distribution of the differences between days inside and outside the challenge periods for the participants.

As the top panel in Figure 2(a) shows, there was a significantly higher number of days with recorded trips in the challenge period. This may imply that some participants turn off SMART when there is no challenge. Figure 2(b) shows a clearly skewed distribution with a peak located around 0. In most cases, SMART reports more or less the same fraction of days with trips inside and outside of the challenge period. However, the mean is significantly greater than 0 (greater than 0.15). For some participants, trips appear to have been underreported outside the challenge period. To correct for this possible bias, we only included participants who had a smaller than 0.2 average difference on both sides of the distribution. This left us with 1056 participants and 3269 person-months data. The bottom panel of Figure 2(a) shows the results after the correction. The figure shows there are still more days with recorded trips in the challenge periods. The remaining difference may be attributed to the fact that there were more workdays in the challenge periods, and days without trips are relatively rare during workdays. The remaining difference may also represent a real effect of the challenges, namely, that participants might have become encouraged to cycle on days on which they normally would not make a trip.

2.4.4 | Data analysis

To analyse short-term behaviour, we compared the users' travel behaviour inside and outside the challenge period for each month. This analysis was done using descriptive statistics. We compared the behaviours in the groups to gain a better understanding of the influence of the monthly challenges on cycling.

The longitudinal analysis focused on quantifying the traveller's behaviours across the months. To capture long-term behavioural change, we only included users who used SMART for at least six months. All users have continued six months of data; therefore, there is no missing data.

As already explained, we had no 'before' data as users could immediately use all functionalities of the SMART app upon downloading. We used ANOVA tests with repeated measures in the SPSS software package to identify statistically significant differences between time-series averages.

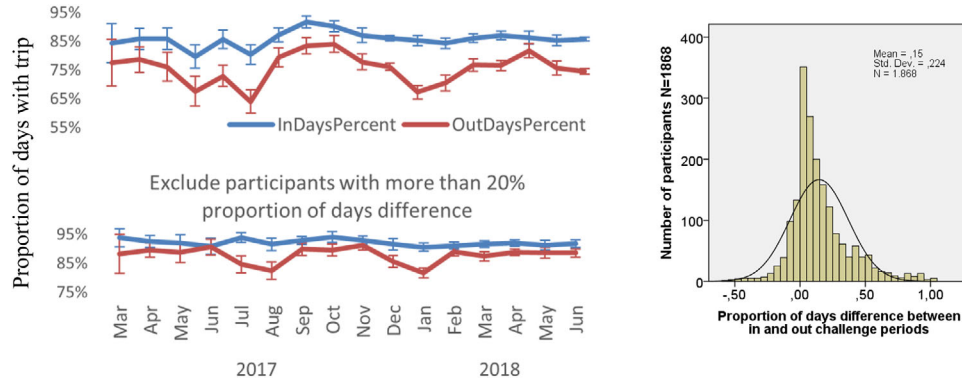


FIGURE 2 Differences between behaviour inside the challenge periods and outside the challenge periods: (a) proportion of days with trip recorded, (b) histogram of proportion difference

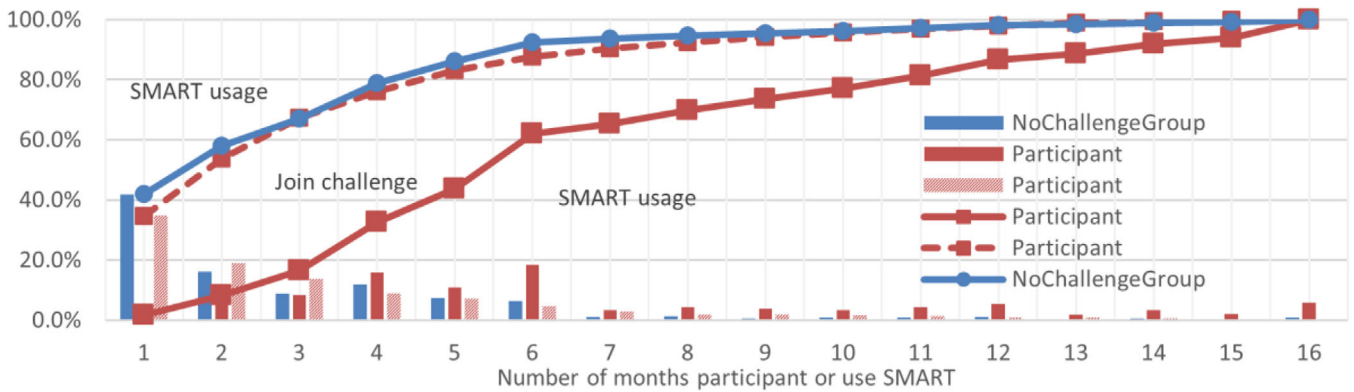


FIGURE 3 The distribution and cumulative user distribution of SMART app usage times or joined the monthly choice challenge times

3 | RESULTS

This section begins with an analysis of the users, followed by a description of the short-term effects of incentives. Finally, long-term behavioural changes are analysed. This is mainly limited to statistically significant results, that is, using a significance level of $p = 0.05$. We mention it explicitly when differences are not statistically significant. The figures include error bars values and indicate two times the standard deviation of the mean.

3.1 | Participants and SMART usage

After data selection, 1057 participants remained in the experiment group and 3657 users in the No Challenge group. Figure 3 shows the distribution over the number of months. Participants on average used SMART over a longer period than users in the No Challenge group. About 40% of the 1057 participants used SMART for six months or longer. However, around 40% of the 1057 participants joined a challenge only once, while 10% joined six times or more. In other words, we could only use a small part of the sample for the long-term analysis. In the whole 18-month period, only 68 participants and 36 users in the No

Challenge group continuously used the SMART app for more than one year. Table 3 shows how long users used the SMART app continuously. For example, we had 526 active participants and 196 users in the No Challenge group who used the app for six continuous months. We used the data in Table 3 to examine long-term trends.

TABLE 3 Descriptive information under different SMART usage conditions and segmentations

Usage	Experiment group	Diff	Easy	Active Participants	Last week
6 months	105	52	53	66	109
7-9 months	57	27	28	39	48
10-12 months	64	26	37	46	52
≥13 months	66	29	36	41	60

Notes: Control group: 87 users who used the app for six months.
 Diff: More than half of the chosen challenges were difficult.
 Easy: More than half of the chosen challenges were easy.
 Active group: Participants who joined the monthly challenge more than three times in the final six months of the study.
 Last week: Participants with trips in the last week of each month.

TABLE 4 Behavioural change for regular and non-regular trips between the inside and outside the challenge periods, with the two sigma error

		Regular			Non-regular			Δ Total
		In	Out	Δ In out	In	Out	Δ In out	Δ In out
Mode share	Bike	70% (0.01)	66% (0.01)	3.4% (0.007)	46% (0.01)	41% (0.01)	3.4% (0.008)	3.4% (0.007)
	Car	18% (0.01)	21% (0.01)	-3.1% (0.007)	41% (0.01)	46% (0.01)	-4.9 % (0.008)	-4.8 % (0.006)
Daily distance	Bike	3.76 (0.13)	3.01 (0.11)	0.75 (0.08)	4.59 (0.13)	3.81 (0.10)	0.78 (0.10)	1.54 (0.13)
	Car	1.34 (0.12)	1.39 (0.11)	-0.04 (0.08)	5.04 (0.19)	5.26 (0.17)	-0.22 (0.14)	-0.33 (0.16)

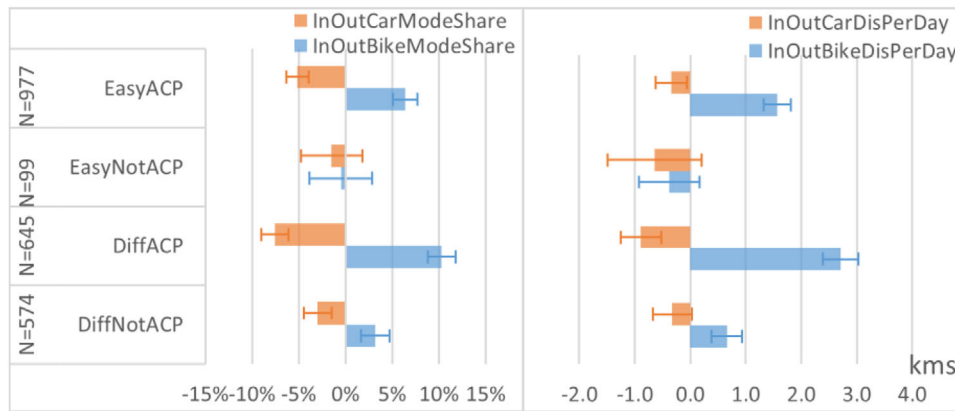


FIGURE 4 Behavioural change inside and outside the challenge periods, with mode share change on the left, and daily distance change on the right, the error bars indicating the two-sigma error

3.2 | The direct effects of monthly challenges

The No Challenge group had an average bike share of 37.8%. The bike share for participants of monthly choice challenges, at 47.4%, is clearly higher. This implies that the monthly choice challenge might attract travellers who are already interested in cycling. Hence, there may be some self-selection among participants. However, it should be noted that for the other SMART users, the observed car mode shares for trips shorter than 20 km are lower than what is found in National Travel Surveys. Even so, the participants still have a high car mode share, which has the potential to get further decreased.

Table 4 shows the travel behaviour changes and allows comparing differences between inside and outside the challenge period. On average, the bike mode share was higher and car share was lower in challenge periods. The daily bike distance was also greater in challenge periods. However, the car distance was not significantly lower.

In Table 4, we distinguish between regular and non-regular trips. Around 30% of the total covered distance was made by regular trips (at least eight trips per month between the same origin and destination), 65% by non-regular trips, and the rest by round trips. Table 4 shows that regular trips had a higher bike share and a lower car share, mostly because trip distances on average are shorter for regular trips. However, behavioural changes were greater for non-regular trips. This suggests that positive incentives to encourage cycling may be more effective for non-regular trips.

In SMART, participants can freely select any monthly challenge. The question is how participants choose and what the effect may be on behavioural change. Therefore, we used challenge difficulty and challenge accomplishment to categorise person-months into four groups: (1) A participant chose a difficult challenge and accomplished the challenge in the corresponding month (*DiffACP*), (2) a participant chose a difficult challenge but did not accomplish the challenge (*DiffNotACP*), (3) a participant chose an easy challenge and accomplished the challenge (*EasyACP*), and (4) a participant chose an easy challenge but did not accomplish the challenge (*EasyNotACP*). We also excluded the first month for each participant because it was not possible to assess the relative difficulty of the challenge based on previous behaviour.

Figure 4 shows the travel behavioural changes for the four groups. More than 90% of the participants who chose easy challenges also accomplished the challenge, while the completion rate was about 50% for participants who chose difficult challenges. On the other hand, the strongest behaviour change occurred for participants who accomplished difficult challenges. This suggests that participants need to choose difficult challenges to obtain large positive effects. However, bike and car shares also increased and decreased, respectively, when participants accomplished easy challenges. The observed overall behaviour change was more or less the same, regardless of whether participants chose easy or difficult challenges. While the change in behaviour was relatively large when participants completed difficult

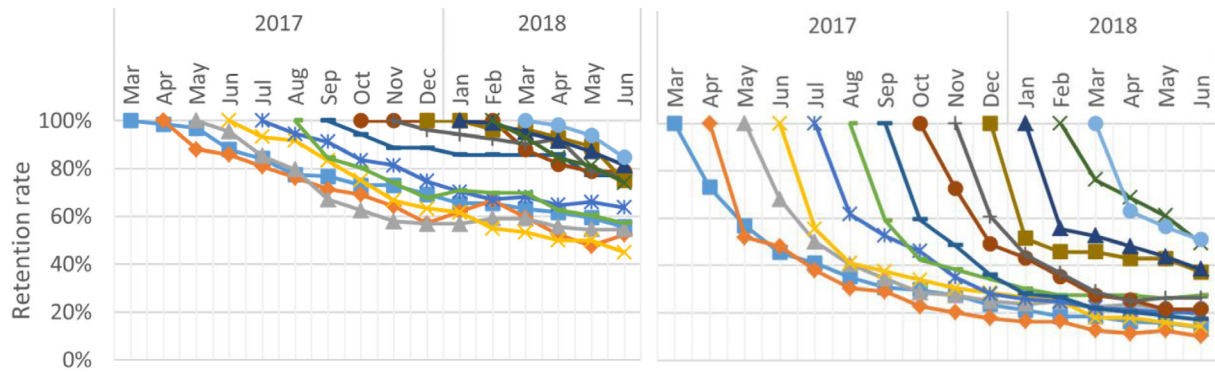


FIGURE 5 User retention: (a) participant group, (b) No Challenge group

TABLE 5 Number of users who continuously used the SMART app in these periods, with challenge participation and total SMART app usage information

Number of users continuously used SMART app during	No Challenge Group	Use SMART ≥ 12 months,	Participants	Joined challenge ≥ 3 times	Use SMART ≥ 12 months, joined ≥ 3 times
June 2017 to June 2018 (13 months)	36	36	68	52	52
One year (12 months)	52	52	89	69	69
Eight months	150	66	261	184	65
Six months	280	63	526	375	71

challenges, difficult challenges were completed much less frequently.

These results were unexpected. Obviously, participants who choose easy challenges do not need to change their behaviour to accomplish the challenge. The fact that they did suggests that the challenges in themselves encourage participants to change their behaviour. Perhaps participants are more aware of their behaviour when they participate in a challenge and therefore more likely to change their behaviour even when they can accomplish the challenge without behaviour change. This would be a positive result and suggests that choice challenges encourage behaviour change in the short term even when participants do not need to change their behaviour to accomplish the challenge.

3.3 | Longitudinal analysis

New users downloaded and started using the SMART app throughout the whole study period. Figure 5 shows the retention rate of users over time after they start using the app. For both participants in choice challenges and users in the No Challenge group, most of the dropouts occurred in the first six to nine months. After this period, retention rates tended to stabilise. The average retention rate was greater than 50% for participants but around 30% or even lower for the No Challenge group. Importantly, users who started using the SMART app after September had a higher retention rate, which appeared to be linked to an update of the SMART app (updated on 13 August 2017).

Table 5 details the number of participants in different stages, that is, those using the app for six, seven to nine, 10 to 12, and more than 12 months (13 or more). We only included the participants who were still using the app in June 2018 (at the end of the study period). Since the retention rates tended to stabilise after six months, Table 5 contains almost all participants who used the app for six months or longer, and we assume that most of the participants in Table 5 have continued using the app afterwards. For the long-term analysis, it, therefore, makes sense to compare these groups in the January to June 2018 period because they more or less represent the same type of participants but who are in a different stage of usage.

In Figure 6, we show the modal shares for the different groups of participants (right panel) and the control group. The control group consisted of users who did not participate in the challenges (No Challenge group) but still may have been influenced by other features of the app over a longer period of time. To exclude those potential effects and enable a fair comparison between the groups of participants, the control group only includes users who started using the app in January 2018. Figure 6 reveals similar seasonal effects for all groups of participants and the control group. Due to increasingly favourable weather conditions, bike shares increased during the first six months, while car shares declined. As mentioned earlier, participants in challenges tended to have a higher bike share and lower car share than users in the control group as Figure 6 shows. The difference is about 15% points. We assume that this is mainly due to self-selection, but we cannot exclude the possibility that there was also some increase in bike share (and decrease in car share) when participants started choosing monthly challenges.

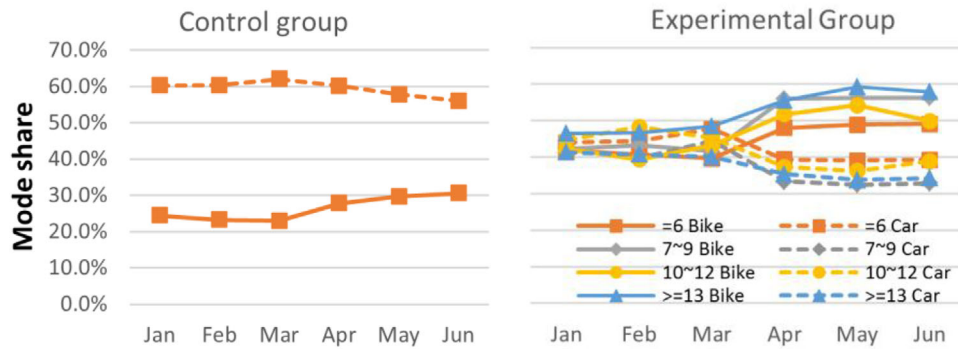


FIGURE 6 Control and experiment groups last six months of travel behaviour under app usage conditions

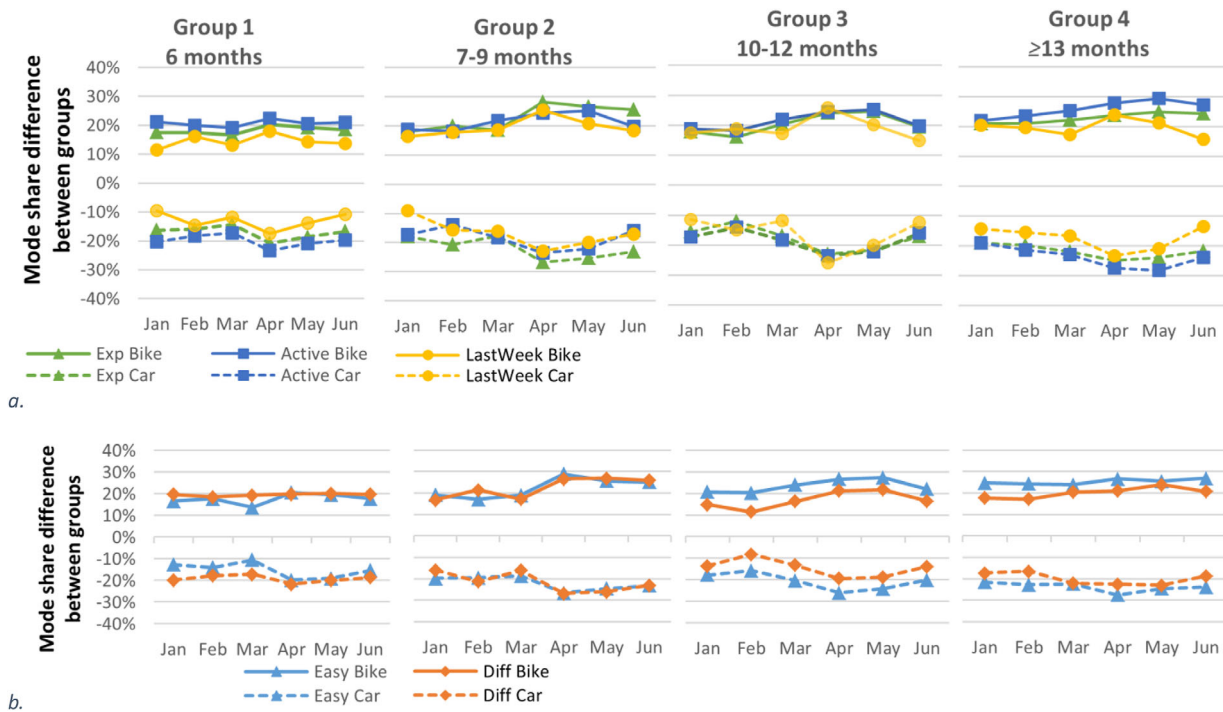


FIGURE 7 Long-term analysis: Comparison of different types of participants with No Challenge group under app usage conditions (a) active participants and non-active participants, (b) easy-challenge, difficult-challenge participants

Unfortunately, we were not able to verify this as we had no ‘before’ measurements.

Figure 7 shows the bike and car mode share difference between the participants and the control group from January to June. We defined groups in the different stages of app usage as follows:

1. Six months (group 1);
2. seven to nine months (group 2);
3. 10 to 12 months (group 3);
4. more than 12 months (group 4).

We compared each group with the control group separately for bike and car share differences. Therefore, we run eight

repeated measures ANOVA for the experimental group with the control group. Moreover, we have four segmentations as experimental subgroups. Thus, there were 40 ANOVA tests in total. Table 6 presents the repeated measures ANOVA results related to Figure 7. The challenge join rate and the share of easy and difficult challenges for each group are also shown in Table 6. Power analysis was operated by the program GLIMMSE as recommend by Guo et al. [55]. The result shows that a sample size of 40 travellers per group, or a total of 80 travellers, would give a power of at least 0.8 for testing the hypothesis of whether there is a time \times intervention interaction. For our study, all ANOVA tests fitted this requirement. Users in all groups had continuously six months of data. Therefore, there is no missing data.

TABLE 6 Repeat ANOVA results

Participants and usage months	ANOVA													
	Avg. bike join	Avg. bike distance mode share	Avg. apbRate	Avg. diff rate	participants versus control groups	Sphericity ϵ	Greenhouse- ϵ	Interaction within subject	between group	Linear contrasts interaction	Quadratic contrasts interaction	Level 4 vs. 1	Level 5 vs. 1	Level 6 vs. 1
Participants	6	3.33	0.45	0.57	Bike share	$V_{isolated}$	$\epsilon = 0.74$	$P < 0.001$	$P < 0.001$			$F(1, 190) = 3.75$ $P = 0.054$ $R = 0.14$		
					Car share	$V_{isolated}$	$\epsilon = 0.70$	$P < 0.001$	$P < 0.001$					
	7-9	4.79	0.49	0.65	Bike share	$V_{isolated}$	$\epsilon = 0.74$	$F(3.66, 519.25) = 6.01$ $P < 0.001$	$P < 0.001$	$F(1, 142) = 9.37$ $P = 0.003$ $R = 0.25$	$F(1, 142) = 14.82$ $P < 0.001$ $R = 0.31$	$F(1, 142) = 8.54$ $P = 0.004$ $R = 0.24$	$F(1, 142) = 5.98$ $P = 0.016$ $R = 0.20$	
					Car share	$V_{isolated}$	$\epsilon = 0.78$	$F(3.85, 546.60) = 3.70$ $P = 0.006$	$P < 0.001$	$F(1, 142) = 4.288$ $P = 0.040$ $R = 0.17$	$F(1, 142) = 9.24$ $P = 0.003$ $R = 0.25$	$F(1, 142) = 5.50$ $P = 0.020$ $R = 0.19$		
	10-12	7.02	0.47	0.65	Bike share	$V_{isolated}$	$\epsilon = 0.84$	$F(4.14, 616.22) = 3.55$ $P = 0.006$	$P < 0.001$	$F(1, 149) = 3.397$ $P = 0.067$ $R = 0.15$	$F(1, 149) = 6.0$ $P = 0.015$ $R = 0.20$	$F(1, 149) = 5.33$ $P = 0.022$ $R = 0.19$		
					Car share	$V_{isolated}$	$\epsilon = 0.73$	$F(3.61, 537.88) = 4.20$ $P = 0.003$	$P < 0.001$	$F(1, 149) = 3.137$ $P = 0.079$ $R = 0.14$	$F(1, 149) = 7.24$ $P = 0.008$ $R = 0.22$	$F(1, 149) = 3.57$ $P = 0.061$ $R = 0.15$		
	>12	8.55	0.49	0.66	Bike share	$V_{isolated}$	$\epsilon = 0.84$	$P < 0.001$	$P < 0.001$		$F(1, 121) = 4.0$ $P = 0.047$ $R = 0.18$			
					Car share	$V_{isolated}$	$\epsilon = 0.75$	$P < 0.001$	$P < 0.001$					
Active	6	4.44	0.47	0.62	Bike share	$V_{isolated}$	$\epsilon = 0.75$	$P < 0.001$	$P < 0.001$			$F(1, 124) = 6.90$ $P = 0.010$ $R = 0.23$	$F(1, 124) = 5.96$ $P = 0.016$ $R = 0.21$	
					Car share	$V_{isolated}$	$\epsilon = 0.69$	$P < 0.001$	$P < 0.001$	$F(1, 124) = 10.07$ $P = 0.002$ $R = 0.27$	$F(1, 124) = 5.04$ $P = 0.027$ $R = 0.20$	$F(1, 124) = 5.60$ $P = 0.020$ $R = 0.21$	$F(1, 124) = 2.99$ $P = 0.086$ $R = 0.15$	
	7-9	6.03	0.50	0.67	Bike share	$V_{isolated}$	$\epsilon = 0.78$	$F(3.79, 469.74) = 5.35$ $P < 0.001$	$P < 0.001$	$F(1, 124) = 4.87$ $P = 0.029$ $R = 0.19$	$F(1, 124) = 4.09$ $P = 0.045$ $R = 0.17$			
					Car share	$V_{isolated}$	$\epsilon = 0.78$	$F(3.83, 475.16) = 3.17$ $P = 0.015$	$P < 0.001$	$F(1, 124) = 4.08$ $P = 0.038$ $R = 0.18$	$F(1, 124) = 4.38$ $P = 0.038$ $R = 0.18$			
	10-12	8.57	0.48	0.68	Bike share	$V_{isolated}$	$\epsilon = 0.87$	$F(4.29, 561.52) = 2.42$ $P = 0.043$	$P < 0.001$	$F(1, 126) = 4.06$ $P = 0.046$ $R = 0.18$	$F(1, 126) = 3.04$ $P = 0.084$ $R = 0.15$	$F(1, 126) = 4.79$ $P = 0.030$ $R = 0.19$	$F(1, 126) = 2.79$ $P = 0.097$ $R = 0.15$	
					Car share	$V_{isolated}$	$\epsilon = 0.79$	$F(3.84, 503.43) = 3.05$ $P = 0.018$	$P < 0.001$	$F(1, 126) = 3.78$ $P = 0.054$ $R = 0.17$	$F(1, 126) = 6.54$ $P = 0.012$ $R = 0.22$			
	>12	11.02	0.52	0.67	Bike share	$V_{isolated}$	$\epsilon = 0.77$	$F(3.78, 476.55) = 2.72$ $P = 0.032$	$P < 0.001$	$F(1, 126) = 3.78$ $P = 0.054$ $R = 0.17$	$F(1, 126) = 6.09$ $P = 0.015$ $R = 0.21$			

(Continues)

TABLE 6 (Continued)

Participants and usage months	ANOVA													
	Avg. join	Avg. bike distance mode share	Avg. apRate	Avg. diff rate	participants versus control groups	Sphericity	Greenhouse-geisser	Interaction within subject	between group	Linear contrasts interaction	Quadratic contrasts interaction	Level 4 vs. 1	Level 5 vs. 1	Level 6 vs. 1
Easy	6	4.11	0.44	0.76	Bike share	Violated	$\epsilon = 0.80$	P < 0.001	P < 0.001	P < 0.001	F(1, 121) = 4.21 P = 0.042 R = 0.18	F(1, 122) = 3.946 P = 0.050 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.22	F(1, 112) = 5.518 P = 0.021 R = 0.22
7-9	5.04	0.49	0.72	0.34	Bike share	Violated	$\epsilon = 0.78$	F(3.75, 423.15) = 4.29 P = 0.003	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22
10-12	7.19	0.50	0.72	0.30	Bike share	Violated	$\epsilon = 0.83$	F(4.09, 498.85) = 2.01 P = 0.091	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22
>12	7.25	0.52	0.73	0.25	Bike share	Violated	$\epsilon = 0.82$	F(3.79, 461.90) = 2.77 P = 0.1029	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22
Diff	6	3.42	0.46	0.48	Bike share	Violated	$\epsilon = 0.72$	P < 0.001	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22
7-9	4.81	0.49	0.63	0.82	Bike share	Violated	$\epsilon = 0.75$	F(3.71, 415.43) = 4.16 P = 0.003	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22
10-12	7.00	0.43	0.56	0.75	Bike share	Violated	$\epsilon = 0.82$	F(3.77, 421.90) = 3.34 P = 0.011	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22
>12	10.4	0.47	0.60	0.64	Bike share	Violated	$\epsilon = 0.72$	F(3.98, 441.51) = 2.90 P = 0.022	P < 0.001	F(1, 113) = 5.84 P = 0.017 R = 0.22	F(1, 113) = 8.53 P = 0.004 R = 0.25	F(1, 122) = 3.946 P = 0.049 R = 0.18	F(1, 112) = 6.759 P = 0.011 R = 0.24	F(1, 112) = 5.518 P = 0.021 R = 0.22

(Continues)

TABLE 6 (Continued)

Participants and usage months	ANOVA													
	Avg. bike distance	Avg. mode share	Avg. acpRate	Avg. diff rate	participants versus control groups	Sphericity	Greenhouse-geisser	Interaction within subject	between group	Linear contrasts interaction	Quadratic contrasts interaction	Level 4 vs. 1	Level 5 vs. 1	Level 6 vs. 1
Last week	6	3.03	0.43	0.58	0.57	Bike share	Violated	$\epsilon = 0.89$	$P < 0.001$					
						Car share	Violated	$\epsilon = 0.92$	$P < 0.001$		$F(1, 184) = 3.56$ $P = 0.061$ $R = 0.14$	$F(1, 184) = 3.28$ $P = 0.072$		
	7-9	4.60	0.48	0.66	0.55	Bike share	Violated	$\epsilon = 0.90$	$P < 0.001$			$F(1, 123) = 4.10$ $P = 0.045$ $R = 0.18$		
						Car share	Violated	$\epsilon = 0.90$	$P < 0.001$		$F(1, 123) = 4.614$ $P = 0.034$ $R = 0.19$	$F(1, 123) = 6.53$ $P = 0.012$	$F(1, 123) = 3.72$ $P = 0.056$ $R = 0.17$	
	10-12	6.35	0.48	0.64	0.49	Bike share	Violated	$\epsilon = 0.91$	$P < 0.001$		$F(1, 127) = 3.72$ $P = 0.056$ $R = 0.17$	$F(1, 127) = 3.032$ $P = 0.084$ $R = 0.15$		
						Car share	Violated	$\epsilon = 0.93$	$P < 0.001$		$F(1, 127) = 6.18$ $P = 0.014$ $R = 0.22$	$F(1, 127) = 7.79$ $P = 0.006$ $R = 0.24$		
	>12	8.67	0.48	0.68	0.38	Bike share	Violated	$\epsilon = 0.93$	$P < 0.001$					
						Car share	Violated	$\epsilon = 0.91$	$P < 0.001$		$F(1, 135) = 4.46$ $P = 0.037$ $R = 0.18$	$F(1, 135) = 3.34$ $P = 0.070$ $R = 0.16$		

The upper panel of the figure displays the trends for all participants per group (green lines). Groups 2 and 3 showed a slow increase in bike share and a decrease in car use over time relative to the control group. We found a significant linear increase ($p < 0.05$ and a middle effect size with $r = 0.15\text{--}0.25$ in corresponding one-way ANOVA with repeated measures in Table 6) in bike share and a decrease in car use after the first month. However, there is no significant trend for groups 1 and 4. This suggests that sustained behaviour change may only take place after several months and tends to be stable after one year. As a result, the bike share fractions are on average about 4% points higher for groups 2, 3 and 4, while car shares are about 3% points lower relative to group 1. However, this difference is only significant for the month of May. In this month, bike shares were 6% points higher for groups 2, 3, and 4 (combined).

Figure 7 presents a long-term analysis. The upper panel of Figure 7 also shows highly active participants who entered the monthly challenge more than three times in the last six months (blue lines). After more than a year of using the app, there was a significant increase and decrease in the bike and car shares, respectively, for the participants in group 4 relative to the control group. Bike shares were also higher in group 4 than for non-active users. This may suggest that sustained behaviour change may be more likely for the most active participants.

To have enough trips in our sample, we took all the data for each month. This means that we also included trips that took place in the challenge periods. When considering long-term trends, it is useful to exclude short-term effects due to the challenges. During the last week of each month, almost all participants are outside of the challenge period. Therefore, we also looked at trips from the final week of each month only (yellow lines). However, the trip sample sizes then became so small that the total variation mostly reflected random variation, making significant differences harder to detect. We, therefore, did not find significant trends over the months. However, for group 1, the participants' bike and car shares appeared to be lower and higher, respectively, during each month's last week. This was expected as this last week was outside the challenge period. Interestingly, this difference became smaller after six months, again suggesting that if sustained behaviour change takes place, this may only happen after several months.

In the lower panel of Figure 7, we distinguish between easy and difficult challenges. The group 'difficult challenges' includes users who entered a difficult challenge more than half of the times. For the long-term behaviour change, Figure 7(b) shows the same trends for all the participants. In the first six months (far left), there is no clear trend while there is an increase over time for groups 2 and 3, and bike shares seem to be higher for groups 2, 3 and 4 (relative to the control group). There also appears to be some difference depending on whether participants chose mainly difficult or mainly easy challenges. The bike distance difference between group 1 and the control group is slightly larger for those who predominantly chose difficult challenges. However, for groups 3 and 4, it is the other way around. Bike shares are higher and car shares are lower for the easy challenges (relative to the control group). In fact, for the difficult challenges, mode shares are almost the same for groups 2, 3,

and 4 relative to group 1. On the other hand, for the choosers of easy challenges, bike shares are on average about 6% points higher and car shares are about 7% points lower for groups 2, 3 and 4 relative to group 1. This result suggests that sustained behaviour is mainly achieved through easy challenges that are also much more often accomplished. However, the results are not conclusive as none of the differences is significant (with a significance level of 0.05) due to the small sample size. However, these differences are still substantial, that is, they are of the same magnitude as the differences between inside and outside the challenge period (slightly above %). The main difference between short-term and long-term behaviour analyses is the sample size. The sample size is still relatively large compared with other studies, and a substantial fraction of the app users participated in at least one monthly challenge. However, although the retention rates are relatively high among the active participants, the number of participants that used the app over a longer period was relatively low as many users joined in a late stage of the study. Moreover, not all of the longer-term users showed sustained behaviour change. Therefore, if substantial sustained behaviour changes occurred, this would only be seen in a small fraction of the users, namely, the ones who downloaded the app and used it for more than six months.

4 | DISCUSSION AND CONCLUSION

In this study, we examined the impact of monthly challenges on travellers' cycling behaviour. In the SMART app, participants were able to choose a cycling-related challenge each month and received points upon completion of a challenge. The app also provided other incentives such as feedback, messages and traffic information.

Over 6000 users downloaded the SMART app with a steady 51% half-year retention rate. About 25% of these users joined the monthly choice challenge, indicating that other interventions, such as self-monitoring and feedback, are also important for attracting users. The users who participated in monthly challenges tended to cycle more, suggesting some form of self-selection among the participants in the monthly challenges. (For the other users, the mode shares were similar to those found in the National Travel Survey). While car use was still substantial among this group of participants, retention rates were relatively high. This shows that there is the potential, at least in theory, to encourage more cycling and reduce car use.

We found no clear evidence for behaviour change among users who never joined any of the monthly challenges, which indicates that for behavioural change, they need to become actively involved.

During the monthly two-week challenge periods, the intervention caused a modal shift from car to bike, suggesting that challenge-and-reward interventions are effective for short-term behavioural change. However, the differences between inside and outside the challenge periods were still relatively small. The distance-based modal shift rate was about 5% points. Interestingly, short-term changes appeared to be greater for non-regular trips than for regular trips. One reason may be that the car mode

share for non-regular trips is greater than for regular trips, which would create more room to change.

We assumed users who have high challenge commitment would have more significant change; however, we observed the overall behaviour change was more or less the same regardless of whether participants chose easy or difficult challenges. These results were unexpected. Obviously, participants who choose easy challenges do not need to change their behaviour to accomplish the challenge. The fact that they did suggests that the challenges in themselves encourage participants to change their behaviour. Perhaps participants are more aware of their behaviour when they participate in a challenge and therefore more likely to change their behaviour even when they can accomplish the challenge without behaviour change. This would be a positive result and suggests that choice challenges encourage behaviour change in the short term even when participants do not need to change their behaviour to accomplish the challenge. In a longer term, users tend to choose easier challenges, and sustained behaviour is more likely achieved through easy challenges that are also much more often accomplished.

There is some indication that the interventions were effective for some long-term behavioural changes. In general, we observed some increase in cycling (relative to the control group) after between 6 and 12 months of usage. After 12 months, the bike share remained slightly larger for participants in the monthly challenges, even more so for highly active challenge participants. This pattern is in accordance with the transtheoretical model, which states that sustained behaviour is only observed after some time (often after one year). However, our sample sizes were too small to arrive at definitive conclusions. If sustained behaviour changes occurred, this may only have been the case for a small fraction of the app's users.

Behaviour change has often been equated with action since action is observable, but decisions to change a habit are preceded by a change in awareness [56, 57] as habitual behaviour is relatively unintentional [58, 59]. The change in awareness before action is often ignored in studies because the monitoring periods when incentives are offered are relatively short. This limits the potential for understanding awareness changes and why people may revert to old habits in the long term. We found that different users may take action over different time periods and that some people may not consciously change behaviour but will return to their old habit after a short period of behaviour change related to an incentive. This suggests that in order to cause conscious changes, experiments need to run long enough to cover all the early stages.

One way to raise awareness is to provide feedback and education; the SMART app does this. In addition, Geller [60] has found that several small incentives can be more effective than one large incentive since this causes individuals to develop internal motivation for behavioural change. In SMART, incentives are relatively small and recurrent. At first sight, it is therefore not obvious how sustained behaviour change could still be increased within the context of our experiment. Also, it is quite hard to control for all these parameters in a real-world environment. However, the interventions had a strong impact on the most active participants and accomplished long-term

behavioural changes. Increasing the variety of challenge types and making the challenges feasible for all types of travellers might increase retention [52], increase the challenge join rate, and lead to more sustained behaviour change.

The limitation of this study is that we only analysed the period during which challenges were provided. We did not have 'before' data and had no information on the behaviour after no more monthly challenges would be provided. Moreover, there is self-selection for SMART users in the experiment group. Yet, the self-selection group still had a significant car mode share that can be decreased further, and this study proves that the behaviours of the self-selection group have the potential to be changed in the long term by the interventions. Moreover, some SMART users were in the control group who did not join the challenge yet but used the app for other functions and can join the challenge later. In line with this, future research could focus on to design apps to attract travellers who do not cycle often or focus on an even longer-term experiment because these users may need longer interventions to change in awareness. In addition, in order to let travellers truly behave as in a real-life environment, the real-world experiment was controlled not to involve privacy issues. As a result, there is no sociodemographic data of travellers.

However, this study does show the realistic potential of using positive interventions via a commercial app in the real world. The SMART app provides larger samples than most studies have so far and makes it possible to monitor behaviour continuously over a more extended period of time. As Prochaska and Velicer [61] found with regards to quitting smoking, the so-called maintenance period lasts from six months to about five years. Therefore, we will monitor SMART users for an even longer period of time in a future study.

In conclusion, our study has shown that monthly challenges can be effective persuasive interventions to accomplish short-term change. The effects on the long term are less clear and suggest that behaviour change is a gradual and cyclical process in line with the transtheoretical theory. VTBC schemes need to pay more attention to the different stages of behaviour change and monitor them accordingly. Results from short-term studies probably overestimate the change that can be sustained.

ACKNOWLEDGEMENTS

This study is part of the EMPOWER project, which is funded by the European Union's Horizon 2020 research and innovation programme. The authors thank Johan Koolwaaij at Mobidot for his valuable help with data provision.

REFERENCES

1. Rijksoverheid: Agenda Bicycle 2017–2020. <https://www.fietsberaad.nl/getmedia/1c52943f-8948-4539-8b0d-4fda6048b1a2/Tour-de-Force-Bicycle-Agenda-2017-2020.pdf.aspx> (2016)
2. Brög, W., et al.: Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transp. Policy* 16(6), 281–292 (2009)
3. Sanjust, B., et al.: An impact assessment of a travel behavior change program: A case study of a light rail service in Cagliari, Italy. *Case Stud. Transp. Policy* 3(1), 12–22 (2014)
4. Ben-Elia, E., Ettema, D.: Changing commuters' behavior using rewards: A study of rush-hour avoidance. *Transp. Res. Part F: Psychol. Behav.* 14(5), 354–368 (2011)

5. Lachapelle, U.: Participation and incentive choice of participants in an early vehicle retirement program in Québec, Canada. *Transp. Res. Rec.* 2375, 8–17 (2013)
6. Eliasson, J., Mattsson, L.-G.: Equity effects of congestion pricing: Quantitative methodology and a case study for Stockholm. *Transp. Res. Part A: Policy Pract.* 40(7), 602–620 (2006)
7. Fogg, B.J.: *Persuasive Technology: Using Computers to Change What we Think and Do*. Elsevier Science, Amsterdam (2002)
8. Froehlich, J., et al.: UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In: *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI 2009)*, Boston, MA, USA, p. 1043 (2009)
9. Broll, G., et al.: Tripzoom: An app to improve your mobility behavior. In: *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM 2012)*, Ulm, Germany, p. 1 (2012)
10. Bie, J., et al.: Move better with tripzoom. *Int. J. Adv. Life Sci.* 4(3), 125–135 (2012)
11. Poslad, S., et al.: Using a smart city IOT to incentivise and target shifts in mobility behavior—Is it a piece of pie?. *Sensors (Basel)* 15(6), 13069–13096 (2015)
12. Mattheiss, E.: PEACOX-Persuasive Advisor for CO2-reducing cross-modal trip planning D7.5 Field Trials II Report (2015)
13. Jylhä, A., et al.: MatkaHupi: A persuasive mobile application for sustainable mobility. In: *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp'13 Adjunct)*, Zurich, Switzerland, pp. 227–230 (2013)
14. Gabrielli, S., Maimone, R.: Digital interventions for sustainable urban mobility: A pilot study. In: *UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing*, Zurich, Switzerland, pp. 119–122 (2013)
15. Jariyasunant, J., et al.: Quantified traveler: Travel feedback meets the cloud to change behavior. *J. Intell. Transp. Syst. Technol. Plann. Oper.* 19(2), 1–16 (2013)
16. Jariyasunant, J., et al.: Overcoming Battery Life Problems of Smartphones When Creating Automated Travel Diaries. University of California Transportation Center, Berkeley, CA, USA (2014)
17. Hu, X., et al.: Behavior insights for an incentive-based active demand management platform. *Int. J. Transp. Sci. Technol.* 4(2), 119–133 (2015)
18. Schrammel, J., et al.: Peacox-Persuasive Advisor for CO2-reducing cross-modal trip planning. In: *1st Workshop on Behaviour Change Support Systems @ Persuasive Technology (BCSS 2013)*, Sydney, Australia (2013)
19. Anagnostopoulou, E., et al.: Persuasive interventions for sustainable travel choices leveraging users' personality and mobility type. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10809, pp. 229–241. Springer Verlag, Berlin (2018)
20. Tsirimpa, A., et al.: A reward-based instrument for promoting multimodality. *Transp. Res. Part F: Psychol. Behav.* 65, 121–140 (2019)
21. Oinas-kukkonen, H., Harjumaa, M.: Persuasive systems design: Key issues, process model, and system features. *Commun. Assoc. Inf. Syst.* 24(1), 28 (2009)
22. Schneider, H., et al.: Understanding the mechanics of persuasive system design: A mixed-method theory-driven analysis of the mobile fitness coach freeletics. In: *34rd Annual Association for Computing Machinery Conference on Human Factors in Computing Systems*, San Jose, CA, USA, pp. 309–320 (2016)
23. Oduor, M., Oinas-Kukkonen, H.: Committing to change: A persuasive systems design analysis of user commitments for a behaviour change support system. *Behav. Inf. Technol.* 1–19 (2019)
24. Brynjarsdóttir, H., et al.: Sustainably unpersuaded: How Persuasion Narrows Our Vision of Sustainability. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, TX, USA (2012)
25. Dastjerdi, A.M., et al.: Use intention of mobility-management travel apps: The role of users goals, technophile attitude and community trust. *Transp. Res. Part A: Policy Pract.* 126(May), 114–135 (2019)
26. Pronello, C., et al.: The effects of the multimodal real time information systems on the travel behaviour. *Transp. Res. Procedia*, 25, 2677–2689 (2017)
27. Castellanos, S.: Delivering modal-shift incentives by using gamification and smartphones: A field study example in Bogota, Colombia. *Case Stud. Transp. Policy.* 4(4), 269–278 (2016)
28. Coombes, E., Jones, A.: Gamification of active travel to school: A pilot evaluation of the Beat the Street physical activity intervention. *Health Place* 39, 62–69 (2016)
29. Baranowski, T., Frankel, L.: Let's get technical! Gaming and technology for weight control and health promotion in children. *Childhood Obesity* 8(1), 34–37 (2012)
30. Berger, M., Platzer, M.: Field evaluation of the smartphone-based travel behaviour data collection app "SmartMo". *Transp. Res. Procedia* 11, 263–279 (2015)
31. Andersson, A., et al.: Promoting sustainable travel behaviour through the use of smartphone applications: A review and development of a conceptual model. *Travel Behav. Soc.* 11, 52–61 (2018)
32. Bowden, H., Hellen, G.: a data driven, segmentation approach to real world travel behaviour change, using incentives and gamification. In: Müller, B., Meyer, G. (eds.) *Towards User-Centric Transport in Europe*, pp. 173–182. Springer International Publishing, Cham (2019)
33. Seaborn, K., Fels, D.I.: Gamification in theory and action: A survey. *J. Hum. Comput. Stud.* 74, 14–31 (2014)
34. Ben-Elia, E., Ettema, D.: Rewarding rush-hour avoidance: A study of commuters' travel behavior. *Transp. Res. Part A Policy Pract.* 45(7), 567–582 (2011)
35. Shaffer, V.A., Arkes, H.R.: Preference reversals in evaluations of cash versus non-cash incentives. *J. Econ. Psychol.* 30(6), 859–872 (2009)
36. Gneezy, U., et al.: When and why incentives (don't) work to modify behavior. *J. Econ. Perspect.* 25(4), 191–210 (2011)
37. Huang, B., et al.: Multinomial logit analysis of the effects of five different app-based incentives to encourage cycling to work. *IET Intel. Transport Syst.* 1–12 (2018)
38. Locke, E.A.: Toward a theory of task motivation and incentives. *Organ. Behav. Hum. Perform.* 3(2), 157–189 (1968)
39. Locke, E.A., Latham, G.P.: Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *Am. Psychol.* 57(9), 705–717 (2002)
40. Baretta, D., et al.: Implementation of the goal-setting components in popular physical activity apps: Review and content analysis. *Digital Health* 5. <https://journals.sagepub.com/doi/10.1177/2055207619862706> (2019)
41. Habibiipour, A., et al.: How to sustain user engagement over time: A research agenda. In: *22nd Americas Conference on Information Systems*, AMCIS, San Deigo, USA (2016)
42. Gadziński, J.: Perspectives of the use of smartphones in travel behaviour studies: Findings from a literature review and a pilot study. *Transp. Res. Part C: Emerging Technol.* 88, 74–86 (2018)
43. Lehto, T., Oinas-Kukkonen, H.: Persuasive features in web-based alcohol and smoking interventions: A systematic review of the literature. *J. Med. Internet Res.* 13(3), e46 (2011)
44. Wathen, C.N., Burkell, J.: Believe it or not: Factors influencing credibility on the Web. *J. Am. Soc. Inf. Sci. Technol.* 53(2), 134–144 (2002)
45. Pornpitakpan, C.: The persuasiveness of source credibility: A critical review of five decades' evidence. *J. Appl. Social Psychol.* 34(2), 243–281 (2004)
46. Sillence, E., et al.: A framework for understanding trust factors in web based health advice. *Int. J. Hum. Comput. Stud.* 64(8), (1071–5819), 697–713 (2006)
47. Hsu, C.P., et al.: How experience-driven community identification generates trust and engagement. *Online Inf. Rev.* 36(1), 72–88 (2012)
48. Steg, L., et al.: Understanding the human dimensions of a sustainable energy transition. *Front. Psychol.* 6, 805 (2015)
49. Mobidot. <http://www.mobidot.nl/en/index.php> (2019). Accessed January 2019
50. Oinas-Kukkonen, H., Harjumaa, M.: Towards deeper understanding of persuasion in software and information systems. In: *First International Conference on Advances in Computer-Human Interaction, ACHI 2008*, Sainte Luce, Martinique, pp. 200–205 (2008)
51. Cialdini, R.B.: *Harnessing the science of persuasion*. Harvard Bus. Rev. 79(9), 72–79 (2001)

52. Dorn, E., Sahinyan, A.: Online trust and commutegreener! Is online trust enough to create stickiness behavior? Dissertation, University of Gothenburg (2010)
53. Thomas, T., et al.: Automatic trip detection with the dutch mobile mobility panel: Towards reliable multiple-week trip registration for large samples. *J. Urban Technol.* 25(2), 143–161 (2018)
54. Harms, L., Kansen, M.: *KiM–Cycling Facts*. Ministry of Infrastructure and Water Management, Netherlands Institute for Transport Policy Analysis (2018)
55. Guo, Y., et al.: Selecting a sample size for studies with repeated measures. *BMC Med. Res. Method.* 13(1), 100 (2013)
56. Triandis, H.C.: Values, attitudes, and interpersonal behavior. *Nebr. Symp. Motiv.* 27, 195–259 (1980)
57. Triandis, H.C.: *Interpersonal Behavior*. Brooks/Cole Pub. Co., Monterey, CA, USA (1977)
58. Ronis, D.L., et al.: Attitudes, decisions and habits as determinants of repeated behaviour. In: Pratkanis, A.R., Breckler, S.J., Greenwald, A.G. (eds.) *Attitude Structure and Function*. Lawrence Erlbaum Associates, Mahwah, NJ, USA (1989)
59. Gärling, T., Garvill, J.: Psychological explanations of participation in everyday activities. *Adv. Psychol.* 96(C), 270–297 (1993)
60. Geller, E.S.: *Keys to Behavior-Based Safety*. ABS Consulting, Government Institutes, Rockville, MD, USA (2001)
61. Prochaska, J.O., Velicer, W.F.: The transtheoretical model of health behavior change. *Am. J. Health. Promot.* 12, 38–48, (1997)

How to cite this article: Huang B, Thomas T, Groenewolt B, Claasen Y, van Berkum E. Effectiveness of incentives offered by mobile phone app to encourage cycling: A long-term study. *IET Intell Transp Syst.* 2021;15:406–422. <https://doi.org/10.1049/itr2.12034>