Preprint
Not Peer-Reviewed

The impact of a gamified intervention on physical activity in real-life conditions: a retrospective analysis of 4800 individuals

Alexandre Mazéas [1,2,3] *, Cyril Forestier [4], Guillaume Harel [3], Martine Duclos [2,5], Aïna Chalabaev [1]

Data, code & supplemental materials:
https://osf.io/scnu7/

Key words
behavior change; mHealth; gamification; intervention; physical activity.

[1]Univ. Grenoble Alpes, SENS, Grenoble, France
[2]National Research Institute for Agriculture, Food and Environment (INRAE), Clermont-Ferrand, France
[3]Kiplin, Nantes, France
[4]Laboratoire Motricité, Interactions, Performance, MIP - UR4334, Nantes Université, France
[5]Department of Sport Medicine and Functional Exploration, University Hospital Clermont-Ferrand, Hospital G. Montpied, Clermont-Ferrand, France

Abstract

Background. Digital interventions integrating gamification features hold promise to promote physical activity (PA). However, results regarding the effectiveness of this type of intervention are heterogeneous.
Objectives. This study aimed to examine the effectiveness of a gamified intervention and its potential moderators in a large sample using real-world data. Specifically, we tested (1) whether a gamified intervention enhanced daily steps during the intervention and follow-up periods compared to baseline, (2) whether this enhancement was higher in participants to the intervention than in nonparticipants, and (3) what participants' characteristics or intervention parameters moderated the effect of the program.
Methods. Data from 4819 individuals who registered for a mHealth Kiplin program between January 1st, 2019, and January 2nd, 2022 were retrospectively analyzed. In this intervention, participants could take part in one or several games where their daily step count was tracked, allowing individuals to play with their overall activity. Nonparticipants are people who registered to the program but did not take part in the intervention and were considered as a control group. PA was assessed via the daily step count of participants. Exposure to the intervention, the intervention content, and participants' characteristics were included in multilevel models to test the study objectives.
Results. Compared with nonparticipants, participants who benefited from the intervention had a significantly greater increase in mean daily steps from baseline during the same period (p <.0001). However, intervention effectiveness depended on participants' initial PA. Whereas the daily steps of participants with <7500 baseline daily steps significantly improved from baseline both during the Kiplin intervention (+3291 daily steps) and during follow-up periods (+945 daily steps), participants with a higher baseline had no improvement or significant decreases in daily steps after the intervention. Age (p <.0001) and exposure (p <.0001) positively moderated the intervention effect.
Discussion. In real-world settings and among a large sample, the Kiplin intervention was significantly effective to increase the daily steps of participants from baseline, during intervention and follow-up periods, and in comparison to nonparticipants. Interestingly, responses to the intervention differed as a function of individuals' initial PA with the existence of a plateau effect. Based on the insights of the self-determination theory, we can assume that the effect of gamification could depend of the initial motivation and activity of participants.

Introduction

Physically inactive individuals are at higher risk of developing non-communicable diseases – such as cardiovascular diseases, cancers, type 2 diabetes mellitus, or obesity – and mental health issues compared to the most active ones.[1] Yet, one-third of the world's population is insufficiently active[2,3] and the trend is downward, with adults performing on average 1000 fewer steps than 2 decades ago[4]. Additionally, it has recently been reported that the global population step count did not return to pre-pandemic levels in the 2 years since the onset of the COVID-19 outbreak.[5] In this context, there is an urgent need to increase physical activity (PA) of individuals in primary, secondary, and tertiary prevention.

Digital behavior change interventions and more particularly gamified services are promising avenues to promote PA. By their omnipresence in our lives, relatively low cost, ability to access content anywhere quickly, and capacity to collect real-time data in natural context, digital means allow for personalized, scalable, wide-reaching, and cost-effective interventions.[6–8] However, there are concerns about the ability of digital programs to engage participants once the novelty wears off or to be effective on any type of audience regardless of their age, sociodemographic, or health status. In this context, gamification strategies introduce an exciting roadmap for addressing these challenges.

Gamification refers to the use of game elements in nongame contexts[9] and allows to transform a routine activity into a more engaging one. The self-determination theory (SDT)[10] is a commonly used theoretical framework for understanding the motivational impact of gamification on behavior. The SDT suggests the existence of different types of motivation that can be pictured on a continuum ranging from lack of motivation to intrinsic motivation (when the behavior is realized for the sake of the activity itself). Although different forms of motivation can initiate the action, people will be more likely to perform the behavior in the long term when their motivation is intrinsic, in other words, when the activity is performed for its inherent pleasure. From this perspective, a gamified intervention that provides fun and playful experiences would feed the intrinsic motivation of participants and would be more correlated with the long-term adherence to PA.

A recent meta-analysis[11] revealed that digital gamified interventions, lasting on average 12 weeks, improved PA by 1600 daily steps on average. Importantly, the results showed that gamified interventions a) appear more effective than digital non-gamified interventions, b) seem appropriate for any type of user regardless of their age or health status, and c) the PA improvement persists in the long term. As a result, gamified interventions are emerging as high-potential behavior change tools to tackle the physical inactivity pandemic. However, these findings obtained from randomized controlled trials do not always reflect what happens in real-life settings.[12] In addition, the effect sizes reported in this meta-analysis were heterogeneous and the authors found high between-study heterogeneity (e.g., $I^2 = 82\%$).

If this heterogeneity may be explained by differences in study quality or diversity of designs in the included studies, the behavior change intervention ontology proposed by Michie et al.[13] argues that heterogeneity in behavioral interventions could also be explained by different variables such as intervention characteristics (e.g., content, delivery), the context (e.g., characteristics of the population targeted such as demographics, setting such as the policy environment or physical location), exposure of participants with the program (e.g., engagement and reach), and the mechanisms of action (the processes by which interventions influence the target behavior). Considering these variables within gamification contexts could provide a useful means to better understand the conditions under which interventions are successful.

The present study investigated these questions based on a retrospective analysis of real-world data collected from a large sample of participants who were proposed a mHealth gamified intervention developed by the company Kiplin. In this one, participants could take part in one or several collective games where their daily step count was tracked, allowing individuals to play with their overall activity. In addition to offering the possibility of direct intervention on people's activity habits in natural context, the capacity of this mobile app to collect, in real-time, a large amount of objective real-world data can be useful to understand the processes and outcomes of behavioral health interventions.[14] More specifically, these data can help make explicit when, where, for whom, and in what state for the participant, the intervention will produce the expected effect, notably thanks to continuous data collection over time. The within-person evolution in daily steps obtained via the app combined with between-person individual factors and intervention parameters is of great interest in this perspective.

Thus, the objectives of this study were to analyze the data collected in order (1) to examine within-individual evolutions of PA before, during, and after the intervention, (2) to test the effectiveness of a gamified program in real-life conditions on PA of participants versus nonparticipants, and (3) to explore the variables that could explain heterogeneity in response to the intervention. Based on previous results on gamification [11], we first hypothesized that PA would increase during and after the gamified program compared to baseline (H1). Second, we hypothesized that this improvement will be greater for participants than for non-participants (i.e., participants who registered on the app but did not complete any game, H2). Finally, we expected that the intervention's characteristics (i.e., type and number of games), the context within the intervention was performed (i.e., population and settings), and the exposure to the intervention (i.e., engagement of participants with the app) will moderate the intervention effect (H3).

Methods

Study design and participants

This study retrospectively analyzed data from adult participants who had registered for a Kiplin program and had given consent for their data to be collected. To be included, participants must be 18 years old or older, have registered on the app between January 1st, 2019, and January 2nd, 2022, and logged daily steps (measured via their smartphone or an activity monitor) on a time frame of at least 90 days with less than 20% of missing daily observations. Of the 134,040 individuals who registered on the Kiplin app on this timespan, 4819 met the eligibility criteria.

Non-wear days were defined as days with fewer than 1000 steps and considered as missing observations – as previous research suggested that daily step values less than 1000 may not represent full data capture.[15,16] Days before the first day of the first game were considered as 'baseline' (14 days ± 42.9, median ± standard deviation), the period between the first day of the first game and the last day of the last game as 'intervention period' (19 days ± 31.2), and the days after the last day of the last game as 'follow-up' (90 days ± 22.8). We restricted the follow-up periods to 90 days post-intervention (i.e., 3 months).

Participants could receive the Kiplin intervention a) in the context of their work (i.e., primary prevention with employees), b) in a senior program (i.e., primary prevention with volunteer retirees), or c) as part of their chronic disease care (i.e., patients mainly treated for obesity or cancer). In all the aforementioned conditions, the program was paid not by the participant but by their employer or health care center.

Some participants registered for the program, created their account, but did not take part in the intervention (i.e., did not completed any game). These individuals were considered "nonparticipants" and were used as a control group (as proposed in previous research [17]). Similarly, the baseline period of these nonparticipants corresponds to the days prior to the date they were supposed to start the intervention period. See Figure 1 for the study flow chart.

The study was approved by local Ethics Committee (IRB00013412, CHU de Clermont Ferrand IRB #1, IRB number 2022-CF063) with compliance to the French policy of individual data protection.

The Kiplin intervention

Participants benefited from one or several PA games (depending on the program) lasting approximately 14 days each. If several games were proposed, these games followed each other in an interval of fewer than 60 days. Details on the games' content have been reported previously.[18] In short, the Kiplin games are collective games where the daily step count performed by each participant is converted into points and allow each team to progress on a map (Fig. 2A), unlock clues in the context of an investigation (Fig. 2B), advance on a boardgame (Fig. 2C), or win a race (Fig. 2D). These games include a multitude of gamification mechanisms such as points, trophies, leaderboards, a chat, challenges, narratives – mechanics that are closely linked to proven behavior change techniques.[19]

Variables

The variables of interest were selected on the basis of the behavior change intervention ontology of Michie et al.[13] and included (1) the longitudinal evolution of daily steps, (2) the exposure of each participant to the intervention, (3) the intervention parameters, and (4) the context (participants' characteristics and settings), as these variables are likely to influence the intervention effect. Table 1 specifies the measures of interest and their operationalization.

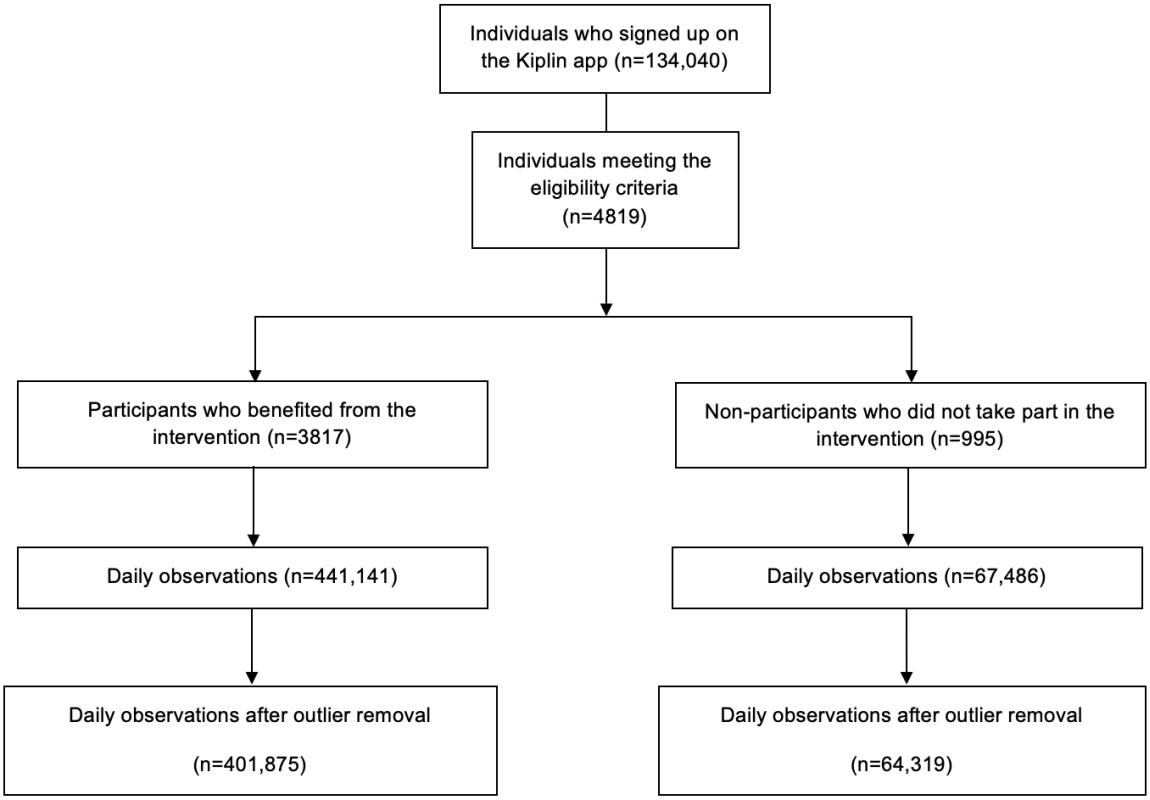

Figure 1. Study flow chart and screening of the databases.

Table 1. Operationalization of the variables.

Outcome	Operationalization
Primary outcome – target behavior (dependent variable)	
Daily step count	PA was assessed via the daily step count, measured with the smartphone or activity monitor of the participant. The daily step count is a trusted proxy for PA.[20] During onboarding, participants were asked to connect to their tracking device (e.g., Apple Health, Google Fit, Fitbit, Garmin) for synchronization of their step count data. In this way, the daily step count of the participants was automatically synchronized on the Kiplin app and the app could retrieve the daily step count for the previous 15 days.
Intervention (content and delivery) and mechanisms of action	
Type of game	Participants could play 4 types of games (i.e., challenge, adventure, boardgame, mission).
Exposure	
Compliance ratio	The engagement of participants with the app was computed as the compliance ratio representing the number of days with a login during the game period divided by the duration of the game periods. This variable allows measuring the frequency of the engagement with the service.[21]
Number of games played	The total number of games played during the intervention period.
Context (population and setting)	
Self-reported age and gender	Filled out by participants when they registered on the app.
Population	Employees, seniors, or patients (treated for obesity, or cancer).
Cofounders	
Season	The season (winter, spring, summer, autumn) when the data step was logged, was controlled as the season can influence PA.[22]

Type of device	The type of device used to assess daily step count (i.e., Android or iOS smartphones, Garmin, Withings, Polar, Fitbit, or Tomtom wearables) was controlled as smartphone apps and wearable devices differ in accuracy and precision.[23]
Lockdown	The study period was characterized by the COVID-19 pandemic and 3 lockdowns were set up in France to limit the spread of the outbreak. As these periods had a strong influence on PA of individuals,[24] we controlled the lockdown periods in our analyses.

Statistical analyses

We calculated the step count increase by subtracting the baseline average daily step count from the average daily step count during the intervention or follow-up periods for each participant and then computed the relative change in %. Mixed-effects models were used to 1) analyze within-person evolution across time (i.e., changes in daily steps between baseline, intervention, and follow-up periods), and across participants and nonparticipants, and 2) examine the associations between intervention parameters, exposure to the intervention, participants' characteristics and settings, and the daily steps evolution. This statistical approach controls for the nested structure of the data (i.e., multiple observations nested within participants), does not require an equal number of observations from all participants,[25] and separates between-person from within-person variance, providing unbiased estimates of the parameters.[26,27]

First, an unconditional model (i.e., with no predictor) was estimated for each variable to calculate intra-class correlations (ICC) and estimate the amount of variance at the between and within-individual levels, which allowed us to determine whether conducting multilevel models was relevant or not. Then, a model that allowed random slope over time (i.e., model with random intercept and random slope) was compared to the null model (i.e., with only random intercept) using an ANOVA, to evaluate whether the less parsimonious model explain a significantly higher portion of the variance of the outcome, compared to the unconditional model.[28,29] Third, between-level predictors and confounding variables were added to another model (Model 1)*[1] and compared to the previous models. Finally, intervention characteristics as well as their interactions with the phases (i.e., baseline/intervention/follow-up) of the study were added in a final model excluding nonparticipants (Model 2)*[2].

Model fit was assessed via the Bayesian Information Criterion (BIC) and −2-log-likehood (−2LL).[30] All models were performed using the *lmerTest* package in the R software.[31] An estimate of the effect size was reported using the marginal and conditional pseudo R^2. When the interaction terms turned significant, contrasts analyses were computed using the *emmeans* package.[32] Models' reliability (estimated with residual analyses) and outliers detection were performed using the *Performance* package.[33] Sensitivity analyses were conducted using all data (including data before outlier imputation) and are available in supplementary materials.

The data and code for the statistical analyses used in the present study are available on Open Science Framework (https://osf.io/scnu7/).

*[1] The equation for the Model was the following: $Y_{ij} = (\beta_0 + \gamma_{0i} + \theta_{0j}) + (\beta_1 + \theta_{1j})\ Time_j + \beta_2\ Phase_j + \beta_3\ Age_j + \beta_4\ Sex_j + \beta_5\ Population_j + \beta_6\ Season_j + \beta_7\ Captor_j + \beta_8\ Baseline\ PA_j + \beta_9\ Lockdown_j + \beta_{10}\ Condition_j \times Phase_j + \varepsilon_{ij}$ where β_0 to β_{10} are the fixed effect coefficients, θ_{0j} and θ_{1j} are the random effect for the participant j (one random intercept and one random slope), γ_{0i} is the random effect for the Time i (random intercept), and ε_{ij} is the error term.

*[2] The equation for the Model was the following: $Y_{ij} = (\beta_0 + \gamma_{0i} + \theta_{0j}) + (\beta_1 + \theta_{1j})\ Time_j + \beta_2\ Phase_j + \beta_3\ Age_j \times Phase_j + \beta_4\ Sex_j + \beta_5\ Population_j \times Phase_j + \beta_6\ Season_j + \beta_7\ Captor_j + \beta_8\ Baseline\ PA_j \times Phase_j + \beta_9\ Lockdown_j + \beta_{10}\ Compliance\ ratio_j \times Phase_j + \beta_{11}\ Number\ of\ games\ played_j \times Phase_j + \beta_{12}\ Type\ of\ Game_j + \varepsilon_{ij}$ where β_0 to β_{12} are the fixed effect coefficients, θ_{0j} and θ_{1j} are the random effect for the participant j (one random intercept and one random slope), γ_{0i} is the random effect for the Time i (random intercept), and ε_{ij} is the error term.

Figure 2. Screenshots of the Kiplin games.

Results

Descriptive Results

Descriptive results are presented in Table 2. The final sample included 4812 adults (mean age = 42.7 ± 11.5 years; 60% women). Sociodemographic variables and baseline physical activity were not descriptively different between participants and nonparticipants. Participants wore an activity monitor measuring their daily step count for an average of 113 days (range = 90 – 686 days). 34,922 daily steps observations were missing on a total of 544,449, which is equivalent to 6.4% of missing data on the full dataset.

Is the gamified program effective to promote PA? (H1)

During the intervention period, participants increased their daily steps by 2619 steps per day on average (+55.6%), compared to the baseline period, and by 317 steps per day on average during the follow-up period (+13.8%), compared to the baseline. In comparison, the daily step count of the control group remained more or less stable throughout the same timeframe with a mean increase of 151 daily steps compared to baseline (+7.5%).

Overall, contrast analyses of the model for the intervention participants (Model 2, Table 3) revealed a negative effect of the intervention on the daily step count during the intervention phase compared to baseline activity (b = -0.09, 95 CI [-0.14; -0.05], p <.0001) and no significant effect (b = 0.01, 95 CI [-0.05; 0.06], p = 0.79) during follow-up periods compared to baseline. However, the patterns were different when participants were stratified by baseline PA. Participants with lower baseline daily steps (<5000 steps per day or 5001-7500 steps per day) showed a significant increase of their daily steps during the intervention and the follow-up, both compared to the baseline (respectively b = 0.25, 95 CI [0.22; 0.28], p <.0001 and b = 0.12, 95 CI [0.09; 0.15], p <.0001). Participants with initial values between 7501 and 10000 steps did not have a significant increase in their daily steps during the intervention (b = 0.00, 95 CI [-0.05; 0.05], p = 0.99) nor during the follow-up period (b = -0.01, 95 CI [-0.04; 0.02], p = 0.44), compared to baseline. Participants who performed more than 10000 baseline steps had significant deteriorations during the intervention (b = -0.13, 95 CI [-0.19; -0.08], p <.0001) and follow-up (b = -0.06, 95 CI [-0.10; -0.03], p = 0.0001). These trends are depicted in Figure 3 and in Table 4. Results were similar in sensitivity analyses that used data without outlier imputation, except for participants with initial daily step count between 7501 and 10000 daily steps who observed significant improvements during and after the intervention (Table S1 and S2 in supplementary materials).

Is the intervention effect greater for participants compared to nonparticipants? (H2)

In Model 1 (Table 3), participants who received the intervention from Kiplin had a significantly greater increase in mean daily steps between baseline and the intervention period, compared with nonparticipants (b = 0.54, 95%CI [0.52; 0.58], p <.0001). The results were similar in sensitivity analyses (Table S2 in supplementary materials).

Table 2. Descriptive statistics.

	Participants	**Nonparticipants**
Participants	N=3817	N=995
Sociodemographics		
Age, mean (SD)	43.2 (11.08)	41.0 (12.81)
Female (%)	2313 (62)	510 (53)
Employees (%)	3,526 (92)	978 (98)
Patients (%)	194 (5)	17 (2)
Seniors (%)	97 (2)	-
Exposure		
Compliance ratio, mean (SD)	0.84 (0.23)	0
Games played, mean (SD)	1.28 (0.9)	0
In-game days, mean (SD)	22.06 (16.24)	0
Observations in each type of game		
Adventure (% of all days)	21,316 (33)	-
Boardgame (% of all days)	4,093 (6)	-
Challenge (% of all days)	32,801 (50)	-
Mission (% of all days)	6,915 (11)	-
Type of device used		
Android smartphone (%)	1076 (28)	286 (29)
iOS smartphone (%)	810 (21)	533 (54)
Fitbit (%)	750 (20)	52 (5)
Garmin (%)	1,071 (28)	109 (11)
Polar (%)	5 (0.1)	-
Tomtom (%)	3 (0.08)	-
Withings (%)	90 (2)	9 (1)
Observations in each season		
Winter (% of all days)	110,517 (24)	17,451(24)
Spring (% of all days)	94,961 (20)	21,162 (30)
Summer (% of all days)	129,039 (27)	8,804 (12)
Fall (% of all days)	138,429 (29)	24,086 (34)
Observations in each lockdown		
1st lockdown – spring 2020 (% of all days)	10,872 (3)	925 (1)
2nd lockdown – fall 2020 (% of all days)	32,298 (8)	4,110 (6)
3rd lockdown – spring 2021 (% of all days)	23,435 (6)	1,757 (2)

.

Table 3. Mixed effect models' estimates.

	Unconditional Model			Model 1			Model 2		
	b [95 CI]	SE	*p*	*b* [95 CI]	SE	*p*	*b* [95 CI]	SE	*p*
Intercept	-0.04 [-0.06; -0.02]	0.01	**<.001**	-0.70 [-0.78; -0.62]	0.04	**<.001**	-0.34 [-0.39; -0.29]	0.03	**<.001**
Fixed Effects									
Time	-0.10 [-0.11; -0.09]	0.01	**<.001**	0.11 [0.05; 0.16]	0.03	**<.001**	-0.05 [-0.06; -0.04]	0.01	**<.001**
Phases of the study									
Baseline				Reference			Reference		
Intervention				0.01 [-0.02; 0.04]	0.01	0.441	0.31 [0.28; 0.34]	0.01	**<.001**
Follow-up				-			0.29 [0.27; 0.32]	0.01	**<.001**
Participants' profile									
Age				0.01 [-0.00; 0.02]	0.01	0.117	-0.00 [-0.02; 0.01]	0.01	0.523
Sex female				Reference			Reference		
Sex male				-0.01 [-0.03; 0.01]	0.01	0.397	0.02 [-0.00; 0.05]	0.01	0.104
Condition									
Control				Reference			-		
Kiplin				-0.03 [0.07; 0.09]	0.02	0.079	-		
Population									
Workers				Reference			Reference		
Seniors				-0.05 [-0.12; 0.03]	0.04	0.213	0.13 [0.05; 0.21]	0.04	**0.002**
Obese patients				-0.12 [-0.21; -0.02]	0.05	**0.017**	-0.01 [-0.14; 0.12]	0.06	0.864
T2DM patients				-0.37 [-0.62; -0.12]	0.13	**0.003**	-0.18 [-0.44; 0.09]	0.13	0.189
Cancer patients				0.02 [-0.05; 0.09]	0.04	0.532	0.07 [-0.01; 0.15]	0.04	0.076
Other patients				-0.11 [-0.21; -0.01]	0.05	**0.031**	0.08 [-0.05; 0.21]	0.07	0.241
Type of game									
Adventure							Reference		
Boardgame							-0.17 [-0.20; -0.14]	0.02	**<.001**
Challenge							-0.08 [-0.10; -0.06]	0.01	**<.001**
Mission							-0.17 [-0.20; -0.14]	0.01	**<.001**
Exposure									
Observance ratio							0.02 [0.00; 0.04]	0.01	**0.021**

Number of games					0.02 [-0.00; 0.04]	0.01	0.064
Season							
Winter		Reference			Reference		
Spring		0.04 [0.03; 0.06]	0.01	**<.001**	0.04 [0.02; 0.05]	0.01	**<.001**
Summer		0.11 [0.09; 0.13]	0.01	**<.001**	0.06 [0.05; 0.08]	0.01	**<.001**
Fall		0.07 [0.05; 0.09]	0.01	**<.001**	0.02 [0.01; 0.03]	0.01	**<.001**
Type of device							
Android		Reference			Reference		
iOS		-0.05 [-0.08; -0.03]	0.01	**<.001**	0.02 [-0.01; 0.05]	0.01	0.128
Fitbit		0.29 [0.26; 0.33]	0.02	**<.001**	0.50 [0.47; 0.53]	0.02	**<.001**
Garmin		0.17 [0.14; 0.20]	0.02	**<.001**	0.33 [0.30; 0.36]	0.01	**<.001**
Polar		-0.94 [-2.23; -0.35]	0.66	0.154	0.11[-0.72; 0.95]	0.43	0.788
Tomtom		0.75 [0.35; 1.15]	0.21	**<.001**	0.81 [0.45; 1.18]	0.19	**<.001**
Withings		0.05 [-0.03; 0.12]	0.04	0.199	0.11 [0.05; 0.18]	0.03	**0.001**
Lockdown							
Periods without restrictions		Reference			Reference		
1st lockdown		-0.09 [-0.21; -0.01]	0.02	**<.001**	0.18 [-0.20; -0.15]	0.01	**<.001**
2nd lockdown		0.00 [-0.02; 0.03]	0.01	0.795	0.00 [-0.02; 0.02]	0.01	0.989
3rd lockdown		0.01 [-0.01; 0.03]	0.01	0.502	0.02 [0.00; 0.03]	0.01	**0.034**
Random Effects							
Level 1 intercept variance	0.70	0.67			0.67		
Level 2 intercept variance	0.32	1.88			0.13		
-2*log (lh)	-515610.3	-236638.1			-449086		
Akaike Information Criteria	1031232.6	473334.3			898288.4		
Marginal R^2 / Conditional R^2	0.01 / 0.38	0.071 / 0.86			0.225 / 0.39		

What are the moderators of the intervention effect? (H3)

The Model 2 estimates are displayed in Table 2. The variables under consideration explained 39% of the variance in daily steps. In this model, we tested the hypothesized interactions, to investigate predictors associated with the efficiency of the intervention (Table S4 in supplementary materials). Contrast analyses were conducted on significant interactions and revealed that the age (b = 0.05, p <.0001) and the compliance ratio (b = 0.37, p <.0001) were positively associated with the change in daily steps between baseline and intervention. Specifically, the older the age, the more regularly the individuals played and the more effective the intervention was. On the other hand, the number of games played by participants was negatively associated with this change (b = -0.02, p = 0.02). In other words, the longer the intervention and the higher the number of games, the less effective the intervention. For categorical outcomes, contrast analyses revealed differences in the intervention effect between the different populations. Compared to employees, cancer patients (b = -0.18, p <.0001), and seniors (b = -0.19, p <.0001) observed a significantly weaker effect of the intervention in comparison to baseline PA. There was no significant difference between employees and obese patients (b = -0.07, p = 0.13). All the results of these analyses are available in supplementary materials.

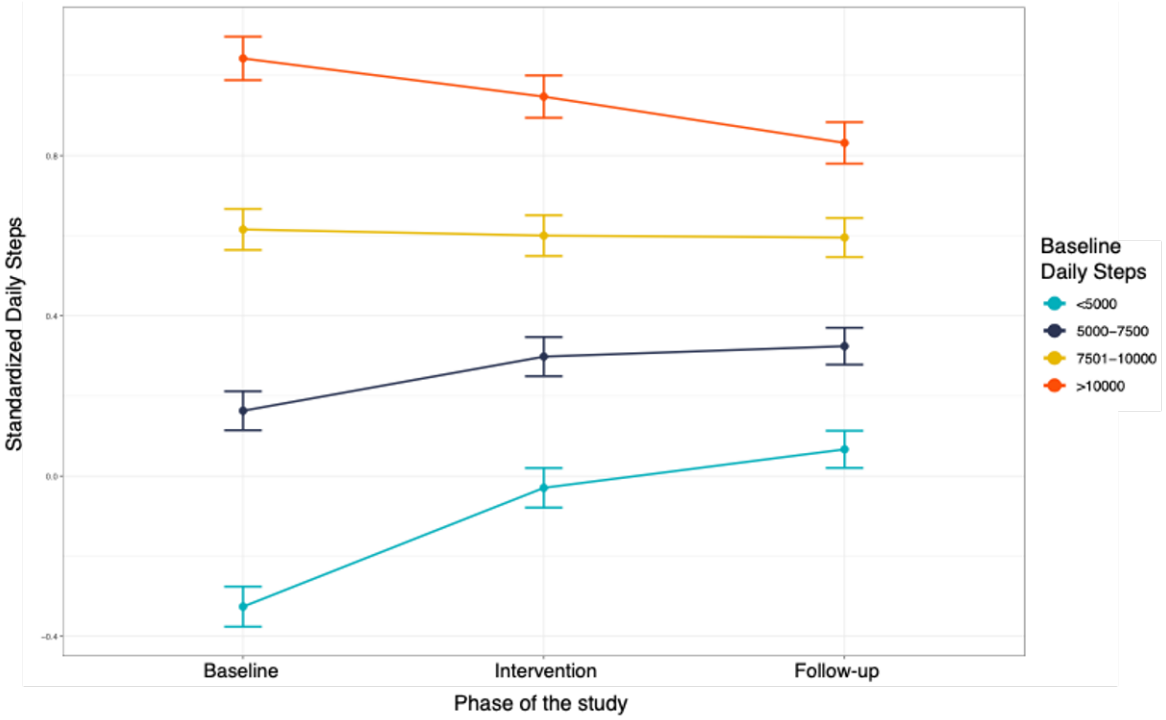

Figure 3. Changes in daily steps throughout the study phases for participants who received a Kiplin program, stratified by baseline daily steps

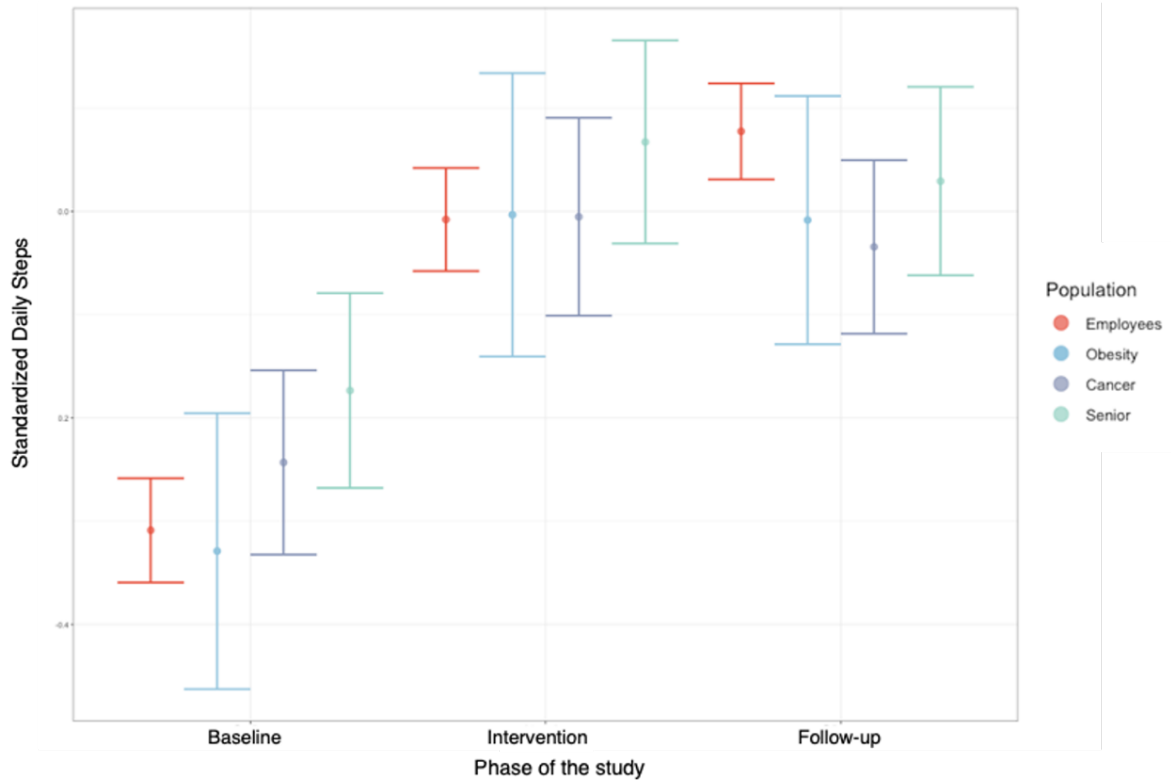

Figure 4. Changes in daily steps across the study phases for the different populations who received a Kiplin program.

Table 4. Description of the mean daily step count during baseline, intervention, and follow-up periods, changes, and relative changes from baseline in function of participants' baseline daily step count.

	<5000	**5000-7500**	**7501-10000**	**>10000**
Baseline daily step count	3671	6096	8818	10111
Intervention daily step count	7490	8855	10301	11388
Follow-up daily step count	5119	6534	7971	9424
Change from baseline during the intervention	3820	2762	2187	1309
Change from baseline during follow-up	1459	431	-156	–697
Relative change during intervention	+118.8 %	+47.2 %	+28.8 %	+16.9 %
Relative change during follow-up	+49.5 %	+8.2 %	-1 %	–4.3 %

Discussion

This study revealed that participants benefiting from the Kiplin games significantly increased PA compared with nonparticipants during the same period. Interestingly, the intervention effect depended on the baseline PA of individuals. Participants with lower baseline steps (<7500 steps per day) significantly improved their PA both during the intervention (between +34% and +76%) and follow-up periods (between +10% and +33%) whereas participants with more than 7500 steps had no significant change or significant decreases. These results suggest that a gamified program is more efficient among inactive individuals than active ones, with the existence of a plateau effect. They also confirm recent findings[11,34] and the ability of gamified interventions to improve PA both during and after the end of the program and in real-life settings[35] – at least for the more inactive individuals. This effectiveness is particularly interesting considering that current behavioral interventions struggle to change PA in the long haul.[36]

The SDT can be a useful framework to explain these contrasting results between initially active and inactive participants. Whereas gamification strategies could enhance the intrinsic motivation of inactive participants as suggested by a previous study,[37] the use of rewards on already motivated people could undermine their intrinsic motivation. Known as the overjustification effect,[38] this phenomenon suggests that if people receive rewards for doing an activity that they used to enjoy, they are likely to discount the internal reason, and thus become less intrinsically motivated than before getting the rewards. This could explain why the same intervention had positive effects on inactive participants who performed more daily steps after the end of the intervention (i.e., during follow-up periods) compared to already active ones who observed significant decreases after the intervention compared to their baseline PA.

The results of this study also stressed that older age may not be incompatible with gamified interventions. Indeed, the intervention effectiveness was moderated by the age of the individual and gamification was more efficient among older individuals, compared to younger ones. These findings are in line with a previous study[39] which found that older users had a greater degree of use of the gamification features. The authors proposed the explanation that older adults pay generally more attention to their health and thus have a stronger intention to engage in a health program. From another perspective and in light of the gamification strategies embedded in the Kiplin intervention, these results could also be explained by the fact that these strategies are accessible – inspired by traditional board game rules and mechanics widely known in the general population – and thus may be more attractive for older populations. Indeed, the most engaging game mechanics may diverge between youth and other populations,[40] and we can expect that younger populations may prefer more complex game mechanics and need more novelty during the intervention to stay interested by the service.

Regarding the effects of the gamified intervention according to the characteristics of the population, a stronger effect was found for programs among employees and obese patients. If these results must be taken with caution considering the high variability observed in patients or senior participants, these findings suggest that gamified interventions are suitable for both primary and tertiary prevention, as suggested by previous work.[11]

The findings also revealed several insights that could help to improve future intervention design. First, exposure to the content is essential for the gamified intervention to be effective. It is interesting as gamification has often been assimilated into a self-fulfilling process permitting automatic engagement of participants. These results are consistent with previous findings demonstrating that higher use of gamification features was associated with greater intervention effectiveness.[39,41] If gamification can ultimately increase program engagement, developers need first to design their apps to be as attractive as possible and optimize retention.

Second, the total number of games played was negatively associated with the intervention effect, suggesting that a shorter intervention could be more beneficial for behavior change. These results are in line with previous research[11,42] suggesting that users benefit more from digital interventions shorter than 3 months. It also suggests a « dose-response » relationship in inverted U shape, with an optimal “middle” to find. Nevertheless, it is necessary to take into consideration the fact that Kiplin programs of more than one game are built in such a way as to decimate several doses at regular intervals. Periods without games were therefore considered in the intervention phases and could explain why, overall, the shorter games were more efficient. More refined analyses of the intervention effect over time will be necessary in the future.

Third, the daily step count of participants was significantly higher in the adventure and the challenge. These two games share the characteristic of being more competitive with a stronger emphasis on leaderboards than the two other games more focused on collaboration. In this vein, Patel et al.[43] observed that the competitive version of their gamified intervention outperformed the collaborative and supportive arms. Moreover, various studies demonstrated that leaderboards are a particularly successful gamification mechanic.[39,44]

Strengths and limitations

This study has several strengths, including the large number of participants included, the intensive objective PA measurement in real-life conditions, and the longer baseline and follow-up duration compared with most trials on gamification that typically incorporate measurement bursts dispersed across time.[11]

However, several limitations should be considered. First, this study was observational and not a randomized controlled trial. Thus, we cannot establish the causality of the intervention's effect on outcome improvement. The nonparticipants are not a true control group. If they did not receive the intervention, it may be for underlying motivational reasons that could impact their PA.

Second, intervention lengths differed between participants. Third, if multilevel models are useful for describing trends in PA behavior change over time, they are limited in their capacity to assess precise fluctuations patterns of non-stationarity behavior such as PA[45] across time and time series analyses could be used in future longitudinal studies to precisely describe patterns of change.

Finally, the compliance ratio used in this study as a proxy for engagement tends to oversimplify the exposure of participants to the service. Complementary measures of engagement (e.g., using the number of logins, time spent per login, and the number of components accessed) will need to be conducted to draw the longitudinal impact of the engagement of the participants on the intervention effect.

Conclusion

In this study in which we retrospectively analyzed the daily step count of 4800 individuals in real-life conditions, participants who benefited from the Kiplin gamified intervention had a significantly greater increase in mean daily steps from baseline than nonparticipants. Responses to the intervention were significantly different as a function of individuals' initial PA. Whereas participants with less than 7500 baseline daily steps had significant improvements both during the intervention and follow-up periods with +3291 daily steps during the program and +945 after the intervention on average, the intervention had no effect on participants with initial values >7500. The motivational effect of gamification could therefore depend on the initial PA and motivational profile of the participants. The age of participants and their engagement with the app were positively and significantly associated with the intervention effect while the number of games played was negatively associated with it. The results of this study suggest that gamification is effective to promote PA in inactive populations with short and medium-term effects. To our knowledge, this study is the first to examine the longitudinal effect of a gamified program outside the context of a trial, with real-world data. The results of this study are therefore highly generalizable and confirm the interest of gamification in both primary and tertiary prevention, and in a broad spectrum of ages.

Additional Information

Contributions

Conceptualization of the study: AM; data curation: AM, GH; investigation: AM; methodology and statistics: AM, CF; software: AM; writing – original draft: AM; writing – review & editing: all the authors.

Declaration of interests

AC, CF, and MD declare that they have no competing interests. AM's PhD grant is funded by the French National Association for Research and Technology (ANRT) and Kiplin. GH is employed by Kiplin.

Funding

The work of AM is supported by an ANRT grant (Cifre PhD Thesis) and by the company Kiplin. The funders had no input in the design of the study and no influence on the interpretation or publication of the study results.

Data & Code Accessibility

The anonymized data used in this study and the R code are available on the Open Science Framework (https://osf.io/scnu7/).

References

1. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. Lancet. 2012. doi:10.1016/S0140-6736(12)61031-9

2. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Heal. 2018;6: e1077–e1086. doi:10.1016/S2214-109X(18)30357-7

3. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, et al. Global physical activity levels: Surveillance progress, pitfalls, and prospects. The Lancet. 2012. doi:10.1016/S0140-6736(12)60646-1

4. Conger SA, Toth LP, Cretsinger C, Raustorp A, Mitáš J, Inoue S, et al. Time Trends in Physical Activity Using Wearable Devices: A Systematic Review and Meta-analysis of Studies from 1995 to 2017. Med Sci Sports Exerc. 2022;54: 288–298. doi:10.1249/MSS.0000000000002794

5. Tison GH, Barrios J, Avram R, Kuhar P, Bostjancic B, Marcus GM, et al. Worldwide physical activity trends since COVID-19 onset. Lancet Glob Heal. 2022;10: e1381–e1382. doi:10.1016/s2214-109x(22)00361-8

6. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. J Behav Med. 2019;42: 67–83. doi:10.1007/S10865-018-9966-Z

7. Walsh JC, Groarke JM. Integrating Behavioral Science with Mobile (mHealth) Technology to Optimize Health Behavior Change Interventions. Eur Psychol. 2019;24: 38–48. doi:10.1027/1016-9040/a000351

8. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016;18(11)e287 https//www.jmir.org/2016/11/e287. 2016;18: e5692. doi:10.2196/JMIR.5692

9. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: Defining "gamification." Proc 15th Int Acad MindTrek Conf Envisioning Futur Media Environ MindTrek 2011. 2011; 9–15. doi:10.1145/2181037.2181040

10. Deci EL, Ryan RM. Handbook of self-determination research. 2002.

11. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the Effectiveness of Gamification on Physical Activity: Systematic Review and Meta-analysis of Randomized Controlled Trials. J Med Internet Res. 2022;24: e26779. doi:10.2196/26779

12. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. Crit Care Med. 2010;38. doi:10.1097/CCM.0B013E3181F208AC

13. Michie S, Thomas J, Johnston M, Aonghusa P Mac, Shawe-Taylor J, Kelly MP, et al. The Human Behaviour-Change Project: Harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. Implement Sci. 2017;12: 1–12. doi:10.1186/S13012-017-0641-5/TABLES/3

14. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing Models and Theories for Digital Behavior Change Interventions. Am J Prev Med. 2016;51: 825–832. doi:10.1016/j.amepre.2016.06.013

15. Bassett DR, Wyatt HR, Thompson H, Peters JC, Hill JO. Pedometer-measured physical activity and health behaviors in U.S. adults. Med Sci Sports Exerc. 2010;42: 1819–1825. doi:10.1249/MSS.0B013E3181DC2E54

16. Kang M, Rowe DA, Barreira T V., Robinson TS, Mahar MT. Individual information-centered approach for handling physical activity missing data. Res Q Exerc Sport. 2009;80: 131–137. doi:10.1080/02701367.2009.10599546

17. Wang G, Yang M, Hong M, Krauss J, Bailey JF. Clinical outcomes one year after a digital musculoskeletal (MSK) program: an observational, longitudinal study with nonparticipant comparison group. BMC Musculoskelet Disord. 2022;23: 1–14. doi:10.1186/S12891-022-05188-X/TABLES/3

18. Mazéas A, Chalabaev A, Blond M, Pereira B, Duclos M. Digital intervention promoting physical activity among obese people (DIPPAO) randomised controlled trial: study protocol. BMJ Open. 2022;12: e058015. doi:10.1136/BMJOPEN-2021-058015

19. Cugelman B. Gamification: What it is and why it matters to digital health behavior change developers. J Med Internet Res. 2013;15: 1–6. doi:10.2196/games.3139

20. Saint-Maurice PF, Troiano RP, Bassett DR, Graubard BI, Carlson SA, Shiroma EJ, et al. Association of Daily Step Count and Step Intensity with Mortality among US Adults. JAMA - J Am Med Assoc. 2020;323: 1151–1160. doi:10.1001/jama.2020.1382

21. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring Engagement in eHealth and mHealth Behavior Change Interventions: Viewpoint of Methodologies. J Med Internet Res 2018;20(11)e292 https//www.jmir.org/2018/11/e292. 2018;20: e9397. doi:10.2196/JMIR.9397

22. Turrisi TB, Bittel KM, West AB, Hojjatinia S, Hojjatinia S, Mama SK, et al. movement behaviors : a scoping review from 2006 to 2020. 2021; 1–26.

23. Piccinini F, Martinelli G, Carbonaro A. Accuracy of Mobile Applications versus Wearable Devices in Long-Term Step Measurements. Sensors 2020, Vol 20, Page 6293. 2020;20: 6293. doi:10.3390/S20216293

24. Genin PM, Lambert C, Larras B, Pereira B, Toussaint JF, Baker JS, et al. How Did the COVID-19 Confinement Period Affect Our Physical Activity Level and Sedentary Behaviors? Methodology and First Results From the French National ONAPS Survey. J Phys Act Heal. 2021;18: 296–303. doi:10.1123/JPAH.2020-0449

25. Sonnentag S. Dynamics of Well-Being. Annu Rev Organ Psychol Organ Behav. 2015;2: 261–293. doi:10.1146/ANNUREV-ORGPSYCH-032414-111347

26. Boisgontier MP, Cheval B. The anova to mixed model transition. Neurosci Biobehav Rev. 2016;68: 1004–1005. doi:10.1016/J.NEUBIOREV.2016.05.034

27. Judd CM, Westfall J, Kenny DA. Experiments with More Than One Random Factor: Designs, Analytic Models, and Statistical Power. Annu Rev Psychol. 2017;68: 601–625. doi:10.1146/ANNUREV-PSYCH-122414-033702

28. Bliese P. Multilevel Modeling in R (2.7) A Brief Introduction to R, the multilevel package and the nlme package. 2022.

29. Finch WH, Bolin JE. Multilevel modeling using Mplus. Multilevel Model Using Mplus. 2017; 1–321. doi:10.1201/9781315165882/MULTILEVEL-MODELING-USING-MPLUS-HOLMES-FINCH-JOCELYN-BOLIN

30. Bollen KA, Harden JJ, Ray S, Zavisca J. BIC and Alternative Bayesian Information Criteria in the Selection of Structural Equation Models. Struct Equ Model a Multidiscip J. 2014;21: 1–19. doi:10.1080/10705511.2014.856691

31. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Softw. 2017;82: 1–26. doi:10.18637/JSS.V082.I13

32. Lenth R, Singmann H, Love J, Buerkner P, Herve M. Emmeans: Estimated marginal means, aka least-squares means. R Packag version. 2018;1: 3. doi:10.1080/00031305.1980.10483031

33. Lüdecke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. J Open Source Softw. 2021;6. doi:10.21105/joss.03139

34. Patel MS, Small DS, Harrison JD, Hilbert V, Fortunato MP, Oon AL, et al. Effect of Behaviorally

Designed Gamification with Social Incentives on Lifestyle Modification among Adults with Uncontrolled Diabetes: A Randomized Clinical Trial. JAMA Netw Open. 2021;4: 1–15. doi:10.1001/jamanetworkopen.2021.10255

35. de Buisonjé DR, Brosig F, Breeman LD, Bloom EL, Reijnders T, Janssen VR, et al. Put your money where your feet are: The real-world effects of StepBet gamified deposit contracts for physical activity. Internet Interv. 2023;31: 100610. doi:10.1016/J.INVENT.2023.100610

36. Dunton GF, Rothman AJ, Leventhal AM, Intille SS. How intensive longitudinal data can stimulate advances in health behavior maintenance theories and interventions. Transl Behav Med. 2021;11: 281–286. doi:10.1093/TBM/IBZ165

37. Höchsmann C, Infanger D, Klenk C, Königstein K, Walz SP, Schmidt-Trucksäss A. Effectiveness of a behavior change technique–based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes: Randomized controlled trial. J Med Internet Res. 2019. doi:10.2196/11444

38. Ryan R, Deci E. Cognitive Evaluation Theory, Part I The Effects of rewards, feedback, and other external events on intrinsic motivation. In: Guilford, editor. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. New York; 2017. pp. 123–157.

39. Maher CA, Olds T, Vandelanotte C, Plotnikoff R, Edney SM, Ryan JC, et al. Gamification in a Physical Activity App: What Gamification Features Are Being Used, by Whom, and Does It Make a Difference? Games Health J. 2022;11: 193–199. doi:10.1089/G4H.2021.0207

40. Schwarz AF, Huertas-Delgado FJ, Cardon G, Desmet A. Design Features Associated with User Engagement in Digital Games for Healthy Lifestyle Promotion in Youth: A Systematic Review of Qualitative and Quantitative Studies. Games Health J. 2020;9: 150–163. doi:10.1089/G4H.2019.0058

41. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. PLoS One. 2017;12: e0173403. doi:10.1371/JOURNAL.PONE.0173403

42. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis. J Med Internet Res. 2019;21: e12053. doi:10.2196/12053

43. Patel MS, Small DS, Harrison JD, Fortunato MP, Oon AL, Rareshide CAL, et al. Effectiveness of Behaviorally Designed Gamification Interventions with Social Incentives for Increasing Physical Activity among Overweight and Obese Adults Across the United States: The STEP UP Randomized Clinical Trial. JAMA Intern Med. 2019. doi:10.1001/jamainternmed.2019.3505

44. De Croon R, Geuens J, Verbert K, Vanden Abeele V. A Systematic Review of the Effect of Gamification on Adherence Across Disciplines. Int Conf Human-Computer Interact. 2021;12789 LNCS: 168–184. doi:10.1007/978-3-030-77277-2_14

45. Heino MTJ, Knittle K, Noone C, Hasselman F, Hankonen N. Studying Behaviour Change Mechanisms under Complexity. Behav Sci 2021, Vol 11, Page 77. 2021;11: 77. doi:10.3390/BS11050077

References

1. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. Lancet. 2012. doi:10.1016/S0140-6736(12)61031-9

2. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Heal. 2018;6: e1077–e1086. doi:10.1016/S2214-109X(18)30357-7

3. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, et al. Global physical activity levels: Surveillance progress, pitfalls, and prospects. The Lancet. 2012. doi:10.1016/S0140-6736(12)60646-1

4. Conger SA, Toth LP, Cretsinger C, Raustorp A, Mitáš J, Inoue S, et al. Time Trends in Physical Activity Using Wearable Devices: A Systematic Review and Meta-analysis of Studies from 1995 to 2017. Med Sci Sports Exerc. 2022;54: 288–298. doi:10.1249/MSS.0000000000002794

5. Tison GH, Barrios J, Avram R, Kuhar P, Bostjancic B, Marcus GM, et al. Worldwide physical activity trends since COVID-19 onset. Lancet Glob Heal. 2022;10: e1381–e1382. doi:10.1016/s2214-109x(22)00361-8

6. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. J Behav Med. 2019;42: 67–83. doi:10.1007/S10865-018-9966-Z

7. Walsh JC, Groarke JM. Integrating Behavioral Science with Mobile (mHealth) Technology to Optimize Health Behavior Change Interventions. Eur Psychol. 2019;24: 38–48. doi:10.1027/1016-9040/a000351

8. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016;18(11)e287 https//www.jmir.org/2016/11/e287. 2016;18: e5692. doi:10.2196/JMIR.5692

9. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: Defining "gamification." Proc 15th Int Acad MindTrek Conf Envisioning Futur Media Environ MindTrek 2011. 2011; 9–15. doi:10.1145/2181037.2181040

10. Deci EL, Ryan RM. Handbook of self-determination research. 2002.

11. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the Effectiveness of Gamification on Physical Activity: Systematic Review and Meta-analysis of Randomized Controlled Trials. J Med Internet Res. 2022;24: e26779. doi:10.2196/26779

12. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. Crit Care Med. 2010;38. doi:10.1097/CCM.0B013E3181F208AC

13. Michie S, Thomas J, Johnston M, Aonghusa P Mac, Shawe-Taylor J, Kelly MP, et al. The Human Behaviour-Change Project: Harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. Implement Sci. 2017;12: 1–12. doi:10.1186/S13012-017-0641-5/TABLES/3

14. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing Models and Theories for Digital Behavior Change Interventions. Am J Prev Med. 2016;51: 825–832. doi:10.1016/j.amepre.2016.06.013

15. Bassett DR, Wyatt HR, Thompson H, Peters JC, Hill JO. Pedometer-measured physical activity and health behaviors in U.S. adults. Med Sci Sports Exerc. 2010;42: 1819–1825. doi:10.1249/MSS.0B013E3181DC2E54

16. Kang M, Rowe DA, Barreira T V., Robinson TS, Mahar MT. Individual information-centered approach for handling physical activity missing data. Res Q Exerc Sport. 2009;80: 131–137. doi:10.1080/02701367.2009.10599546

17. Wang G, Yang M, Hong M, Krauss J, Bailey JF. Clinical outcomes one year after a digital musculoskeletal (MSK) program: an observational, longitudinal study with nonparticipant comparison group. BMC Musculoskelet Disord. 2022;23: 1–14. doi:10.1186/S12891-022-05188-X/TABLES/3

18. Mazéas A, Chalabaev A, Blond M, Pereira B, Duclos M. Digital intervention promoting physical activity among obese people (DIPPAO) randomised controlled trial: study protocol. BMJ Open. 2022;12: e058015. doi:10.1136/BMJOPEN-2021-058015

19. Cugelman B. Gamification: What it is and why it matters to digital health behavior change developers. J Med Internet Res. 2013;15: 1–6. doi:10.2196/games.3139

20. Saint-Maurice PF, Troiano RP, Bassett DR, Graubard BI, Carlson SA, Shiroma EJ, et al. Association of Daily Step Count and Step Intensity with Mortality among US Adults. JAMA - J Am Med Assoc.

2020;323: 1151–1160. doi:10.1001/jama.2020.1382

21. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring Engagement in eHealth and mHealth Behavior Change Interventions: Viewpoint of Methodologies. J Med Internet Res 2018;20(11)e292 https//www.jmir.org/2018/11/e292. 2018;20: e9397. doi:10.2196/JMIR.9397

22. Turrisi TB, Bittel KM, West AB, Hojjatinia S, Hojjatinia S, Mama SK, et al. movement behaviors : a scoping review from 2006 to 2020. 2021; 1–26.

23. Piccinini F, Martinelli G, Carbonaro A. Accuracy of Mobile Applications versus Wearable Devices in Long-Term Step Measurements. Sensors 2020, Vol 20, Page 6293. 2020;20: 6293. doi:10.3390/S20216293

24. Genin PM, Lambert C, Larras B, Pereira B, Toussaint JF, Baker JS, et al. How Did the COVID-19 Confinement Period Affect Our Physical Activity Level and Sedentary Behaviors? Methodology and First Results From the French National ONAPS Survey. J Phys Act Heal. 2021;18: 296–303. doi:10.1123/JPAH.2020-0449

25. Sonnentag S. Dynamics of Well-Being. Annu Rev Organ Psychol Organ Behav. 2015;2: 261–293. doi:10.1146/ANNUREV-ORGPSYCH-032414-111347

26. Boisgontier MP, Cheval B. The anova to mixed model transition. Neurosci Biobehav Rev. 2016;68: 1004–1005. doi:10.1016/J.NEUBIOREV.2016.05.034

27. Judd CM, Westfall J, Kenny DA. Experiments with More Than One Random Factor: Designs, Analytic Models, and Statistical Power. Annu Rev Psychol. 2017;68: 601–625. doi:10.1146/ANNUREV-PSYCH-122414-033702

28. Bliese P. Multilevel Modeling in R (2.7) A Brief Introduction to R, the multilevel package and the nlme package. 2022.

29. Finch WH, Bolin JE. Multilevel modeling using Mplus. Multilevel Model Using Mplus. 2017; 1–321. doi:10.1201/9781315165882/MULTILEVEL-MODELING-USING-MPLUS-HOLMES-FINCH-JOCELYN-BOLIN

30. Bollen KA, Harden JJ, Ray S, Zavisca J. BIC and Alternative Bayesian Information Criteria in the Selection of Structural Equation Models. Struct Equ Model a Multidiscip J. 2014;21: 1–19. doi:10.1080/10705511.2014.856691

31. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Softw. 2017;82: 1–26. doi:10.18637/JSS.V082.I13

32. Lenth R, Singmann H, Love J, Buerkner P, Herve M. Emmeans: Estimated marginal means, aka least-squares means. R Packag version. 2018;1: 3. doi:10.1080/00031305.1980.10483031

33. Lüdecke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. J Open Source Softw. 2021;6. doi:10.21105/joss.03139

34. Patel MS, Small DS, Harrison JD, Hilbert V, Fortunato MP, Oon AL, et al. Effect of Behaviorally Designed Gamification with Social Incentives on Lifestyle Modification among Adults with Uncontrolled Diabetes: A Randomized Clinical Trial. JAMA Netw Open. 2021;4: 1–15. doi:10.1001/jamanetworkopen.2021.10255

35. de Buisonjé DR, Brosig F, Breeman LD, Bloom EL, Reijnders T, Janssen VR, et al. Put your money where your feet are: The real-world effects of StepBet gamified deposit contracts for physical activity. Internet Interv. 2023;31: 100610. doi:10.1016/J.INVENT.2023.100610

36. Dunton GF, Rothman AJ, Leventhal AM, Intille SS. How intensive longitudinal data can stimulate advances in health behavior maintenance theories and interventions. Transl Behav Med. 2021;11: 281–286. doi:10.1093/TBM/IBZ165

37. Höchsmann C, Infanger D, Klenk C, Königstein K, Walz SP, Schmidt-Trucksäss A. Effectiveness of a behavior change technique–based smartphone game to improve intrinsic motivation and physical

activity adherence in patients with type 2 diabetes: Randomized controlled trial. J Med Internet Res. 2019. doi:10.2196/11444

38. Ryan R, Deci E. Cognitive Evaluation Theory, Part I The Effects of rewards, feedback, and other external events on intrinsic motivation. In: Guilford, editor. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. New York; 2017. pp. 123–157.

39. Maher CA, Olds T, Vandelanotte C, Plotnikoff R, Edney SM, Ryan JC, et al. Gamification in a Physical Activity App: What Gamification Features Are Being Used, by Whom, and Does It Make a Difference? Games Health J. 2022;11: 193–199. doi:10.1089/G4H.2021.0207

40. Schwarz AF, Huertas-Delgado FJ, Cardon G, Desmet A. Design Features Associated with User Engagement in Digital Games for Healthy Lifestyle Promotion in Youth: A Systematic Review of Qualitative and Quantitative Studies. Games Health J. 2020;9: 150–163. doi:10.1089/G4H.2019.0058

41. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. PLoS One. 2017;12: e0173403. doi:10.1371/JOURNAL.PONE.0173403

42. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis. J Med Internet Res. 2019;21: e12053. doi:10.2196/12053

43. Patel MS, Small DS, Harrison JD, Fortunato MP, Oon AL, Rareshide CAL, et al. Effectiveness of Behaviorally Designed Gamification Interventions with Social Incentives for Increasing Physical Activity among Overweight and Obese Adults Across the United States: The STEP UP Randomized Clinical Trial. JAMA Intern Med. 2019. doi:10.1001/jamainternmed.2019.3505

44. De Croon R, Geuens J, Verbert K, Vanden Abeele V. A Systematic Review of the Effect of Gamification on Adherence Across Disciplines. Int Conf Human-Computer Interact. 2021;12789 LNCS: 168–184. doi:10.1007/978-3-030-77277-2_14

45. Heino MTJ, Knittle K, Noone C, Hasselman F, Hankonen N. Studying Behaviour Change Mechanisms under Complexity. Behav Sci 2021, Vol 11, Page 77. 2021;11: 77. doi:10.3390/BS11050077