OPEN

APPLIED SCIENCES

Large-Scale Fandom-based Gamification Intervention to Increase Physical Activity: A Quasi-experimental Study

MASAMITSU KAMADA[1], HANA HAYASHI[1,2], KOICHIRO SHIBA[1,3], MASATAKA TAGURI[4], NAOKI KONDO[5], I-MIN LEE[6,7], and ICHIRO KAWACHI[3]

[1]Department of Health Education and Health Sociology, School of Public Health, Graduate School of Medicine, The University of Tokyo, Tokyo, JAPAN; [2]Down to Earth, Inc., Katori City, Chiba, JAPAN; [3]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA; [4]Department of Data Science, Yokohama City University School of Data Science, Yokohama City, Kanagawa, JAPAN; [5]Department of Social Epidemiology and Global Health, Graduate School of Medicine and School of Public Health, Kyoto University, Kyoto City, Kyoto, JAPAN; [6]Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; and [7]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

ABSTRACT

KAMADA, M., H. HAYASHI, K. SHIBA, M. TAGURI, N. KONDO, I-M. LEE, and I. KAWACHI. Large-Scale Fandom-based Gamification Intervention to Increase Physical Activity: A Quasi-experimental Study. *Med. Sci. Sports Exerc.*, Vol. 54, No. 1, pp. 181–188, 2022. **Purpose:** Gamification, the use of game design elements in nongame contexts, in combination with insights from behavioral economics, has been applied increasingly to behavior change interventions. However, little is known about the effectiveness or scalability of this approach, especially in the long term. We tested a large-scale smartphone-based intervention to encourage physical activity among Japanese baseball fans using gamification techniques that leveraged fandom and interteam competition inherent in sports. **Methods:** A quasi-experimental study was conducted among fans of the Japanese Pacific League. The app, Pa-League Walk, included gamification elements, such as competition between opposing teams' fans based on total daily step counts on game days (>60,000 free downloads since March 2016). We analyzed daily steps of 20,052 app users, supplemented by online survey data of 274 users and 613 matched controls. Difference-in-differences estimators evaluated change in daily steps before and after app installation in users versus matched controls. **Results:** Users' daily steps increased by 574 (95% confidence interval, 83–1064) steps 3 months after installation, compared with controls. The increase was maintained for up to 9 months (559 (99–1018) more steps per day vs baseline), attenuating over a longer follow-up. Positive effect modification was found by high-frequency of the app use ($P < 0.001$) but not by other covariables ($P \geq 0.14$) such as education or income. Days with 10,000-step achievement increased from 24.4% to 27.5% after the additional introduction of incentives (digital player photographs; $P < 0.001$). **Conclusions:** Using existing fandom and solidarity, the gamification app increased physical activity at scale among baseball fans, including people with lower socioeconomic status underrepresented in traditional health programs. **Key Words:** NUDGING, EXERCISE, MHEALTH, BEHAVIOR CHANGE, DISSEMINATION

Despite the health benefits of physical activity (PA), physical inactivity remains highly prevalent globally (1). Population strategies to promote PA are imperative to reduce the burden of physical inactivity (2).

Recently, novel approaches for health promotion have been increasingly applied, such as gamification and behavioral design (derived from behavioral economics) (3–5). Gamification denotes the use of game design elements (e.g., points, levels, social connectivity, competition) to motivate and engage people in nongame contexts (6); previous studies have demonstrated the effectiveness of gamification interventions for increasing PA (3,4). In addition, findings from behavioral economics can help design interventions that anticipate predictable barriers to behavior change and "nudge" people to make healthier choices (3,7,8). A range of intervention designs, including incentives, have been empirically tested for behavior change and categorized in a taxonomy of "choice architecture" techniques (9). However, the long-term effects of gamification and behavioral designs remain unclear (3,4,6,10,11). A systematic review

Address for correspondence: Masamitsu Kamada, M.E.d., Ph.D., Department of Health Education and Health Sociology, School of Public Health, Graduate School of Medicine, The University of Tokyo, S310, Faculty of Medicine Bldg. 3, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; E-mail: kamada@gakushikai.jp.
Submitted for publication April 2021.
Accepted for publication July 2021.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/21/5401-0181/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000002770

showed that the gamification interventions for health promotion that had been tested as of 2015 were of the duration of 4 months or less (12). Although one randomized trial was conducted in 2018–2019 and showed that a longer 6-month behaviorally designed gamification intervention with social incentives increased PA even during the 3-month follow-up after the intervention ceased (4), long-term studies remain understudied. For their long-term effect, it may be necessary to optimally blend and integrate behavioral design elements for a particular application (6).

Furthermore, despite the popularity of PA-promotion apps (software applications for mobile devices), little is known about their effectiveness (11,13,14). High-quality evidence is needed to draw real-world causal inferences, especially for publicly available apps as scalable interventions.

As a key insight for boosting the success of gamification-based behavior change, sport fandom and a sense of connection with a sports team (15,16) have the potential to be powerful draws for habit formation. For example, the Football Fans in Training program successfully attracted male football (soccer) fans to participate in a weight loss program and achieve long-term (3.5 yr) weight loss (17,18). In Japan, baseball is the most popular sport (19). The Pacific League professional baseball league (*Pa-League*) initiated a large-scale smartphone-based project to nudge baseball fans to become more physically active using gamification techniques that leveraged team pride and interteam competition. Considering the global popularity of sports and smartphones, this model may represent a scalable approach to behavior change. However, no study has examined the effectiveness of large-scale implementation (e.g., $n > 10{,}000$) of a fandom-based intervention bridging spectator sports and daily PA. Therefore, we conducted a quasi-experimental study to investigate the effectiveness and long-term effects of this nationwide gamification for promoting PA among users.

METHODS

App and theoretical framework. The Pa-League Walk app became downloadable for free on iPhones and Android phones from March 2016 (>60,000 downloads by December 2019). A marketing company (Pacific League Marketing Corporation) of the baseball league developed and designed the app, based on behavioral science and gamification with baseball fandom as a key insight (Supplemental Digital Content 1, Figure, screenshots of the app, http://links.lww.com/MSS/C409; Supplemental Digital Content 2, Figure, logical framework, http://links.lww.com/MSS/C410). Table 1 describes key features of the app and the relevant behavioral theories and techniques (6,9,20). The app encourages competition between opposing teams' fans based on total daily step counts on baseball game days. An additional reward, "players photo collection," was introduced in October 2016 (beginning of the baseball off-season) as an incentive for achieving 10,000 steps per day—users received a digital photograph of a random player from their favorite team on days when they achieved 10,000 steps. The app could be used anywhere and anytime, including the baseball off-season (October to March). Thus, target users were not limited to stadium visitors. The app does not mention the health benefits of PA to maintain primary focus on connecting fans and enhancing a sense of solidarity.

Study design and measures. This was a quasi-experimental observational study conducted throughout Japan (Fig. 1).

The primary outcome was the change in daily step counts, which has been associated with clinical end points and mortality in previous studies (21,22). The app automatically recorded the number of daily steps using smartphones' built-in accelerometers (additional details in Supplemental Digital Content 3: Methods, http://links.lww.com/MSS/C411). Steps measured by iPhones were moderately accurate in previous studies (23,24). Because Android phones had significantly larger variability

TABLE 1. Key features of the Pa-League Walk app and the relevant behavioral theories and techniques.

Fandom-based gamification	1. Team-based steps competition: (GM no. 6: social connectivity, no. 7: fun and playfulness) Competition between opposing teams' fans based on daily total steps on actual baseball game days, e.g., (Team A fans total = 300M steps) vs (Team B fans total = 200M steps). The app displays each team's ranking on a daily leader board, based on steps achieved by fans.
	2. A push notification for achieving 10,000 steps with a photograph of randomly selected player: (GM no. 3 providing feedback on performance; BCT no. 2.2 feedback on behavior (CT); CA no. A2 feedback)
	3. 10,000 steps reward, "players photo collection": (GM no. 1: goal setting, no. 4: reinforcement; BCT nos. 10.1 and 10.2 material incentive/reward (behavior), nos. 10.4 and 10.5 social incentive/reward (OC); CA no. B4 connect decision to benefit, no. A decision information: anchoring) The users receive a digital photograph of a randomly selected player from their favorite team with a congratulations message on a day when they achieve 10,000 steps.
	4. Badges for lifetime steps (GM no. 1: goal setting, no. 2: capacity to overcome challenges, no. 5: compare progress) Achieved level based on lifetime steps (badges for MVP, the first string, farm team, and sandlot baseball).
	5. Customized screen designs and "baseballized" stats view of steps (GM no. 7: fun and playfulness (theme); BCT no. 8.7: graded tasks (SCogT)) Standby mode with customized designs by the selected favorite team; displaying daily steps in baseball diamond as 2500 steps being on the first base (single hit), 5000 on the second base, 7500 on the third base, and 10,000 back on the home base (home run).
	6. Real-time display of fellow (loyal) contributors (GM no. 5: compare progress; BCT no. 6.2. social comparison; CA no. A3 provide social reference point) Displaying pop-up of avatar and the number of steps of users who just "voted" their steps to the team on app screen (smartphones) of users who are watching the real-time steps competitions.
Other (general) behavior change techniques	7. Feedback on behavior (BCT no. 2.2, CT): Displaying stats and ranking based on steps.
	8. Prompts/cues (BCT no. 7.1, OC): Teaching the users to use the league's schedule of baseball games as cues to walking and joining the steps competition. Occasional push notifications before the major events of the baseball league and the app program (e.g., season opening, special competitions on all-star game days).
	9. Social comparison (BCT no. 6.2 SCompT): Displaying daily steps rankings (self-position) and leader board (prompt identification as a role model) among Facebook and Twitter friends and among the same team's fans.

Gamification strategies, choice architecture techniques (nudging), and other behavior change techniques are categorized based on previous literature (numbers in parentheses above correspond to these articles) (6,9,20).

BCT, behavior change techniques; CA, choice architecture; CT, control theory; GM, gamification; OC, operant conditioning; SCogT, social-cognitive theory; SCompT, social comparison theory.

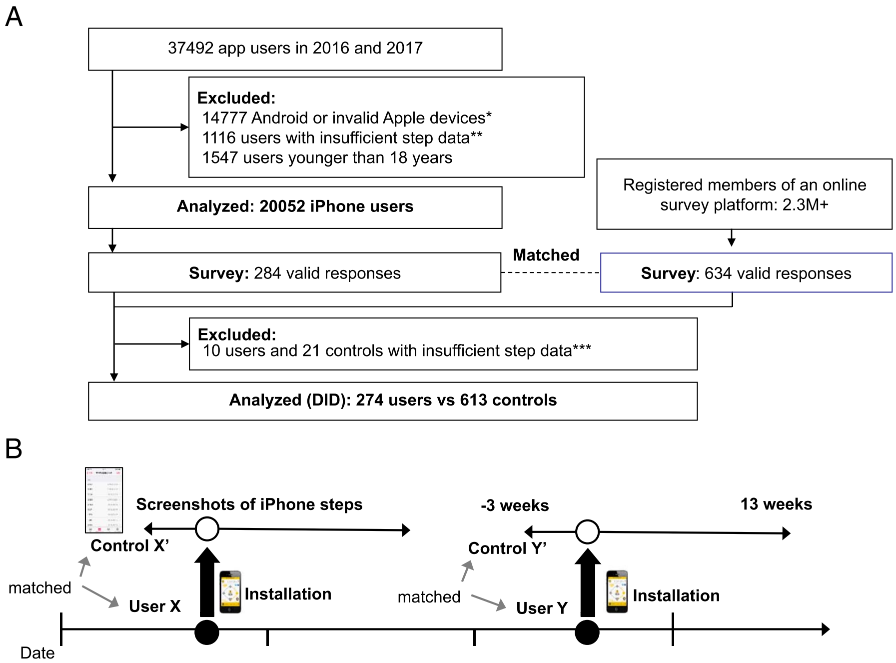


FIGURE 1—Flowchart and matching scheme of the participants. A, Flowchart of the sampling. *Invalid Apple devices include iPad and old iPhone models without motion coprocessor. **Less than 500 steps per day. ***Less than 500 steps per day or <4 valid days each before and after the installation date. B, Matching scheme of the participants. Sex- and age group–matched controls (*X′* and *Y′*) were asked to upload screenshots of their iPhone steps data retrospectively for a period defined by the date of app installation of the corresponding matched users (*X* and *Y*, respectively).

in their acceleration measurements and algorithms and the accuracy of their step measurements is unknown, we analyzed only iPhone data (61% users). All users agreed to the app's privacy policy, which permitted the use of their de-identified data. We used 2016–2017 data from users 18 yr or older, provided by the Pacific League Marketing Corporation. We additionally gathered daily steps before installation of the app (baseline or preperiod data) and daily steps in matched controls during the same period (pre–post control data).

The users declared their sex, age, height, weight, and favorite team among the available six in the league upon initial activation. Area of residence (prefecture) was recorded based on the average GPS location of the phones for those allowing GPS function. Users were invited to an online survey, conducted between December 2017 and January 2018, which gathered information on their daily steps 1 month before app installation (baseline/pre data). Respondents also declared their household annual income, household size, education level, prefecture of residence, frequency of app usage, and frequency of stadium visits. Participants younger than 20 yr at the time of the survey were excluded from the study (participants could be younger than 20 yr, the youngest being 18 yr, on the app installation date).

To obtain matched control data, an online survey was conducted between February and March 2018. Over 2.3 million registered as members of an online survey platform (Rakuten Research, Inc.). Among them, Pa-League baseball fans with iPhones who had not downloaded the app were screened and invited to participate in the survey. Then, each participant (control) was matched by sex and age group to each user, and asked to upload screenshots (photographs) of their iPhone steps data (from a built-in Health app) retrospectively for a period defined by the date of app installation of the corresponding matched users (3 wk before installation through 13 wk after installation; Fig. 1B). Controls answered the same questions as the app users. All respondents gave informed consent. We only considered days with at least 500 steps and participants with at least four valid days each before and after the installation date, leaving 274 users and 613 matched controls (average, 2.2 (range, 1–4) controls per user). There were no missing answers in the survey responses.

As reference information, the distribution of users' stages of change of exercise behavior (leisure-time PA) was obtained from a survey of another sample of users. In September 2019, as a collaborative project between Pacific League Marketing Corporation and Fukuoka City, a survey was conducted among those who newly installed the app and registered the Fukuoka SoftBank Hawks as their supporting team. We received information from the company regarding the distribution of the stages of change at the time of installation among the 591 respondents, which was used as reference information for this study. The question on the stages of change was modified slightly from that in a previous study that tested its reliability and validity (25) and was sent at the first launch of the app as follows: "Please tell us about yourself before you downloaded this app. Please choose one that best describes your exercise habits and thoughts in the past 6 months. Regular exercise here means exercising for 30 min or more per session at least twice a week." Respondents selected either of the following answers: "I did not exercise, and I did not intend to start exercising in the near future (within 6 months)" (precontemplation), "I did not exercise, but I was thinking of starting in the near future (within 6 months)" (contemplation), "I had been recently exercising, but not regularly" (preparation), "I had been exercising regularly, but within 6 months of starting" (action), or "I had been exercising regularly for more than 6 months" (maintenance).

This study used only de-identified data and was approved by the institutional review board of the Harvard T.H. Chan School of Public Health.

Statistical analysis. Baseline characteristics of users and controls were compared using *t*-tests and chi-square tests. Difference-in-differences (DID) analysis using regression models estimated the change in daily steps among app users versus matched controls before and after app installation. In the DID approach, the association between intervention and outcome was estimated by examining the interaction between pre/post and exposed/unexposed variables (26). We used a generalized linear mixed model (GLMM) with fixed effects of sex, age on app installation date (reference date for controls), body mass index (BMI), education, equivalized income calculated by dividing household annual income by the square root of the number of household members, prefecture-level population density, year of app installation, group (app user or control), month (preinstallation; 1, 2, or 3 months after installation), and the group–month interaction; individuals (participants) were included as a random effect. Plausibility of the parallel trends assumption of DID (i.e., postintervention outcome trends between treated and comparison groups would have been the same without the intervention) was evaluated by assessing whether preintervention trends were parallel between groups (additional details in Supplemental Digital Content 3: Methods, http://links.lww.com/MSS/C411).

We also tested the pre–post change in subsamples of users with a similar level of preperiod steps as the matched controls (≤1000 steps per day difference at baseline). Note that the baseline difference in outcome between treatment and control groups is not a threat to validity of the DID analysis, and matching units on pre-period level risks regression-to-the-mean bias (27). Thus, this additional matching was not used for our primary analysis.

In addition, we examined the effect modification of the app by sex, age group (18–39 and 40–68 yr), BMI, education, equivalized income, population density, frequency of app usage, year of app installation, baseball season on the installation date (2016 season, off-season, or 2017 season), frequency of stadium visits, favorite team's rank in steps competitions (top 3 vs bottom 3), and the team's rank in baseball games in the real world (top 3 vs bottom 3). We used the interaction terms between these variables, group, and time (month since installation of the app). We also performed sensitivity analyses: 1) similar GLMM adjusted for baseline average daily steps and 2) adding matched pairs as another random effect and dropping sex and age covariates. To check specific responses of the users to key features of the app, we examined the distribution of daily steps before and after introduction of the 10,000-step-per-day reward system, using 2016 data, and the effect modification by baseball game-days (April–September 2016) when team-based competition on daily steps was conducted via the app (i.e., "home-game" vs "visiting-game" vs "non-game" days). Finally, we examined long-term trends in daily steps over 22 months (i.e., March 2016 to January 2018) after app installation in the DID users ($n = 274$). The duration of the long-term evaluation depended on the app installation date (i.e., the earlier the app was installed, the longer the evaluation period). We compared daily steps after installation with baseline by GLMM, adjusting for the same covariables as the main model and calendar month (seasonality).

Based on the Japanese national average (SD) daily steps of 6642 (4191) (28), a sample size of 277 in each group was necessary to detect 1000-step-per-day change in app users versus controls, using a *t*-test with a two-sided 5% significance level, and a power of 80%. Analyses were carried out using SAS version 9.4 (SAS Institute Inc., Cary, NC).

RESULTS

The sample dataset contained 1.7 million days from 20,052 users (42% female; mean (SD) age, 33.6 (11.6) yr; mean (SD) BMI, 23.0 (3.9) kg·m^{-2}: 24% were overweight (BMI ≥25 kg·m^{-2})) from all 47 prefectures in Japan. The DID sample (274 users vs 613 controls; Table 2) participants were 18–68 yr old (on app installation date), 39% of whom were women. Users had higher BMI and income, were more likely to live in dense areas, and more frequently visited stadiums compared with controls ($P < 0.05$).

Figure 2 shows histograms of daily steps for 14,673 users and 370 controls with valid data in 2016. An additional reward for achieving 10,000 steps per day was introduced in October

TABLE 2. Baseline characteristics of participants for DID analysis ($n = 887$).

	User ($n = 274$)	Control ($n = 613$)	*P*
Age, mean (SD), yr	42.3 (10.7)	41.4 (10.8)	0.27
18–29	40 (14.6)	112 (18.3)	
30–39	63 (23.0)	151 (24.6)	
40–49	93 (33.9)	190 (31.0)	
50–59	68 (24.8)	139 (22.7)	
60–68	10 (3.6)	21 (3.4)	
Sex, female	108 (39.4)	239 (39.0)	0.96
BMI, mean (SD), kg·m^{-2}	23.7 (4.0)	22.6 (3.7)	<0.001
≥25	93 (33.9)	140 (22.8)	0.001
Education			0.92
High school or less	60 (21.9)	142 (23.2)	
Vocational/technical school or junior college	55 (20.1)	122 (19.9)	
4-yr college or higher	159 (58.0)	349 (56.9)	
Equivalized income,[a] Japanese Yen			0.003
≤3,000,000	81 (29.6)	235 (38.3)	
3,000,001–4,999,999	91 (33.2)	215 (35.1)	
≥5,000,000	102 (37.2)	163 (26.6)	
Population density, ≥1000 person·km^{-2}	181 (66.1)	294 (48.0)	<0.001
Frequency of stadium visits, times per year			<0.001
0	14 (5.1)	356 (58.1)	
1–5	92 (33.6)	233 (38.0)	
6–10	55 (20.1)	14 (2.3)	
≥11	113 (41.2)	10 (1.6)	
Year of app installation,[b] 2017	113 (41.2)	243 (39.6)	NA
Frequency of app usage[b]			NA
Less than once a day	83 (30.3)	197 (32.1)	
Once a day	78 (28.5)	173 (28.2)	
More than once a day	113 (41.2)	243 (39.6)	
Valid days of step measurement,[c] median (IQR)	374 (229–543)	105 (95–109)	NA

Values are *n* (%) unless stated otherwise.

[a]Calculated by dividing household annual income by the square root of the number of household members. Approximately 110 yen is equivalent to 1 US dollars.

[b]Control participants were categorized based on the responses of the corresponding matched users.

[c]Days with at least 500 steps. Maximum recorded days were limited to 680 for users and 110 d for controls per study design. The DID analysis used the data 3 wk before installation date through 13 wk after installation both in users and controls (max 112 d).

IQR, interquartile range.

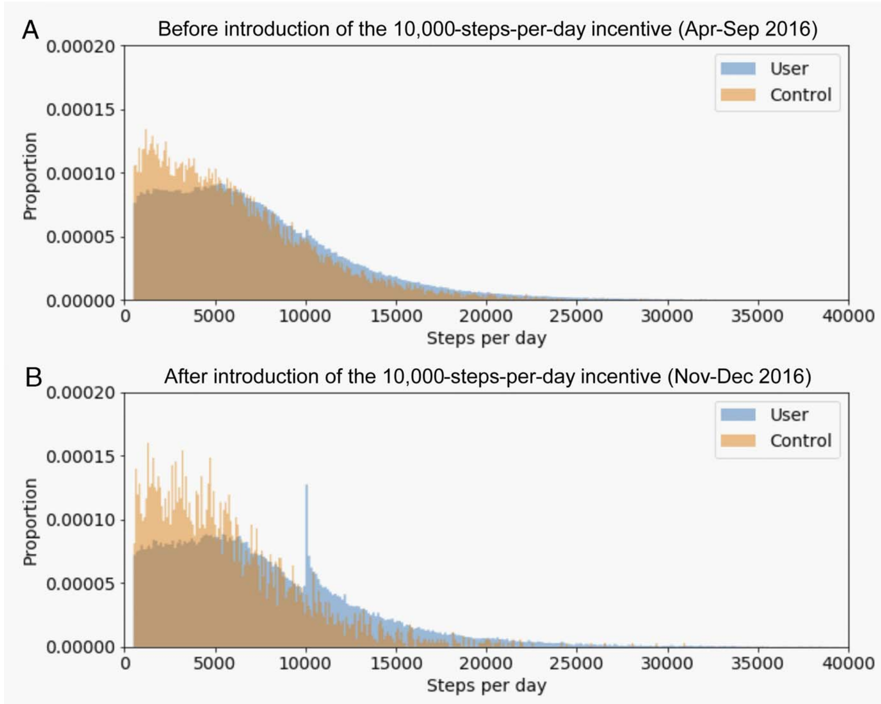


FIGURE 2—Histograms of daily steps for users and controls before (A) and after (B) the introduction of 10,000-step reward system (14,673 users and 370 controls, 2016 season). *Y* axis represents the proportion of the observations (one observation per participant-day) that falls within the bin (relative frequency).

2016. The subsequent distribution of daily steps shows clear bunching around 10,000 steps (Fig. 2B), suggesting that some users began aiming for this level in response to the new reward system. Among users, days with 10,000-step achievement increased from 24.4% to 27.5% after introduction of the incentive ($P < 0.001$), despite the latter being the winter off-season.

Difference-in-differences. Primary DID analysis of 274 users and 613 controls showed that average daily steps for app users increased by 506 (95% confidence interval (CI), 25–986 in month 2) to 574 (83–1064 in month 3) after app installation, versus controls (Fig. 3A). At baseline, users had 2147 (1583–2710) more steps per day than controls. When restricting the sample to 193 participants matched on similar steps (89 users vs 104 controls) before app installation, users' daily steps increased by 1195 (95% CI, 592–1798 in month 3) after installation, compared with controls (Fig. 3B). The restricted users had baseline average daily steps of 6410 (95% CI, 5868–6953), similar to the national average of 6463 steps per day in 2016 (29). In addition, their postperiod average daily steps (7290–7384 steps per day) were similar to that of the entire sample (7164–7464 steps per day in 20,052 users).

Trends in daily steps were similar (parallel) between users and controls before app installation (P (group–time) = 0.23; Supplemental Digital Content 4, Figure, trajectories of average daily steps, http://links.lww.com/MSS/C412). In the control group, the correlation between baseline and pre–post change in daily steps was weak-to-moderate (Pearson's $r = -0.43$, Spearman's

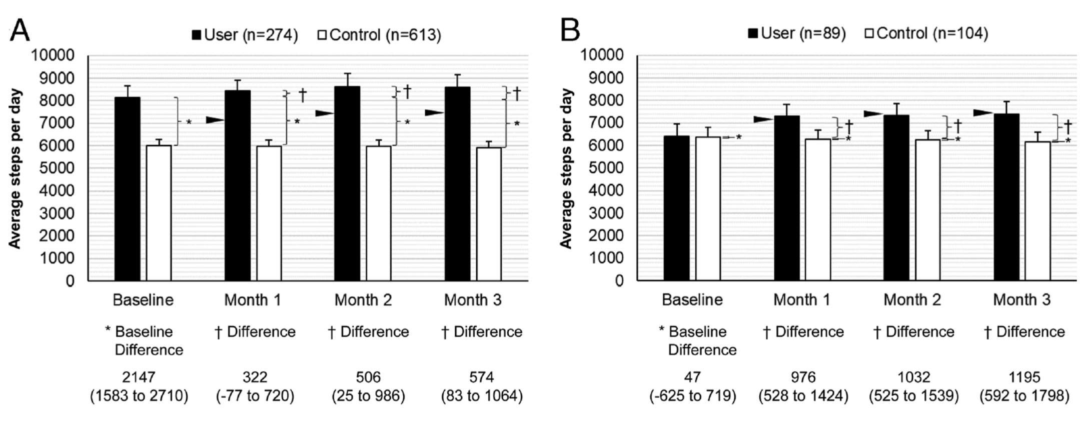


FIGURE 3—Adjusted daily steps before and after installation of the Pa-League Walk app based on the DID model. A, Primary analysis with 274 users vs 613 controls ($n = 887$). B, Subsamples of the matched pairs with <1000-step-per-day difference at baseline ($n = 193$, 89 users vs 104 controls). Error bars are 95% CI, adjusted for potential confounders. Baseline period was 3 wk before app installation. *Triangle tick marks* indicate average steps among all users ($n = 20,052$) for only the period after app installation.

$r = -0.29$; each $P < 0.0001$; Supplemental Digital Content 5, Figure, scatter plot, http://links.lww.com/MSS/C413) with a serial correlation of 0.46 as the autoregressive parameter of the AR(1) correlation structure (27). Considering that control participants with more baseline steps decreased their steps more in the postperiod, the relative increase in daily steps among users (who had more baseline steps, vs controls) can be interpreted as the effect of app usage over the natural trend. Both sensitivity analyses with adjustment of baseline average daily steps and matched pairs as another random effect yielded results similar to the main analysis.

Subgroup analyses. Subgroup analyses showed a significant effect modification by frequency of app use ($P < 0.001$) but not by other covariables ($P \geq 0.14$); the increase in daily steps was greater among those who used the app more frequently (Supplemental Digital Content 6, Table, subgroup analyses, http://links.lww.com/MSS/C414). Greater relative increase in daily steps appeared to be found in those age 40–68 versus 18–39 yr, and among fans with lower educational levels and lower incomes, and those visiting stadiums more frequently. A similar or slightly larger increase in daily steps was found in men than in women. When divided by baseball schedule, the relative increase in daily steps was slightly greater on game days with team-based daily steps competitions, compared with non–game days (interaction $P = 0.07$).

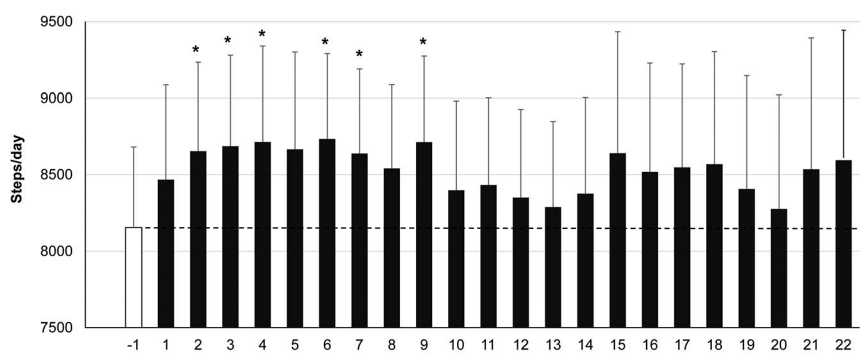
Long-term trend. Figure 4 shows the long-term trend in average daily steps in the DID users ($n = 274$). Expected maximum sample size per month was based on the distribution of installation dates; "follow-up" rate was 94% (44/47) in month 22. Average daily steps peaked 4 months after app installation (559 (95% CI, 72 to 1047) more steps per day compared with baseline). The increase in steps was maintained until month 9 (559 (99 to 1018)), was slightly attenuated, and became unstable over subsequent months (440 (−438 to 1317) in month 22, $P = 0.33$), partly because of smaller sample sizes for estimates of longer maintenance.

Stages of change for exercise. As reference information, a survey of another sample of users ($n = 591$; September 2019) showed that the maintenance stage for exercise behavior was the most common with 28.4% ($n = 168$) at the time of the app installation, followed by precontemplation with 24.5% ($n = 145$). The rest were 23.0% ($n = 136$) in contemplation, 19.0% ($n = 112$) in preparation, and 5.1% ($n = 30$) in action.

DISCUSSION

This quasi-experimental study showed that users' daily steps increased meaningfully (~500 steps per day, equivalent to 0.25 miles, or 5 additional minutes of walking) after app installation. Their average daily steps after app installation (e.g., 7464 steps per day in month 3 in the entire sample) were close to 7500 steps per day of approximate recommended PA level (30). Our causal interpretation is supported by the users' specific responses to unique app functions. Its effectiveness for users from a wide range of socioeconomic backgrounds shows the promise of this fandom-based gamification model as a large-scale behavior change intervention.

The data suggest long-term maintenance (approximately 2 yr) for overall increase in PA. Recent reviews revealed small-to-moderate effects of smartphone apps on PA, mostly of short durations (average ≤3 months) (11,14). Another study showed a short-lived effect of Pokémon GO on daily steps (<6 wk), although there are roughly a billion users worldwide and some people might sustain increased PA (31). The successful long-term effect of the Pa-League Walk might lie in its development based on target insight. The app used fandom, a strong psychological connection with the sport team (i.e., team identification (15,16)), as a key internal motivation for user engagement with an element of competition by proxy (players in the stadium, fans via the app). This may have triggered high user engagement because of a more self-determined form of motivation or regulation (i.e., identified regulation) than, for example, external rewards



No. of Users

Observed	274	272	264	260	253	247	225	215	209	194	183	164	160	157	145	139	128	120	118	106	92	74	44
Expected†	274	274	274	274	274	266	242	235	227	214	203	177	172	166	153	147	139	129	126	114	101	78	47

FIGURE 4—Long-term trend in average daily steps among users ($n = 274$). *$P < 0.05$ compared with baseline (steps before app installation). Bars indicate 95% CI. Adjusted for sex, age, BMI, income, education, population density, calendar month, and year of app installation. †Expected maximum sample size in each month based on the distribution of installation dates. The earlier the app was installed, the longer was the evaluation period. Some users did not have valid data in the first month but did in the following months. An observed sample size less than 274 in the first month does not indicate that some users had no valid data throughout the postperiod.

unrelated to user identities and values (32). Because there are more than 143 baseball games in 6–7 months, the connection between virtual gaming and real-world events (e.g., step competition on game day, communication with fellow fans) might have also reinforced sustained behavior change, accompanied by social (group) cohesion among fans (15). However, it is uncertain whether the improvement in PA would continue once the intervention ceased. A US-based randomized trial showed that PA improvement via competition-based gamification remained during a 3-month follow-up (4). Thus, integrating and optimally blending behavioral design elements, such as competition and reinforcement, may foster a long-term effect (6).

Previous studies have not empirically demonstrated the effectiveness of large-scale implementation of a fandom-based bridge between spectator sports and daily PA. Although health promotion programs by European football clubs successfully maintained PA improvement in male football fans (e.g., 678-step-per-day increase at 12 months in EuroFIT) through their attractive intervention delivery by community coaches in the club stadium, there were few beneficiaries (2,000 or less), and access to the stadium presumably limited geographical range (17,18,33,34). The potential reach of professional sports is far more extensive, with an estimated 27 million baseball fans in Japan (19) and over 700 million soccer fans globally (35). With smartphones' ubiquity, apps can maximize the reach of programs that bridge spectator sports and daily PA.

The app was effective even for people with lower socioeconomic status (SES). People with lower SES are reportedly more likely to be in early stages of change (e.g., precontemplation) regarding PA (36). Conventional public health approaches informing the public about PA's health benefits appeal to health needs. In contrast, the fandom-based gamification intentionally did not use health benefits as a key message. Thus, the app might be perceived as simply an entertainment tool to enjoy spectator sports and is accessible to everyone in any stage of PA change, including lower-SES groups. A subsample survey of the Pa-League Walk app showed that a quarter (24.5%) of users were in a precontemplation stage of exercise behavior before app installation. In addition, the concept of "contribution to team" appeals to belongingness needs, and the "team pride" notion appeals to esteem needs; both are higher-level appeals compared with messages appealing to safety (health) in Maslow's hierarchy of needs (37). Despite the emergence of such a novel approach, disseminating knowledge about the health benefits of PA (outside this app) remains fundamental to public health policy, as part of a systems approach to increasing PA (38).

This study has several strengths. First, we objectively measured PA via a smartphone accelerometer, thereby avoiding recall bias. Second, a quasi-experimental design and analyses supported our causal interpretation. Third, the participants were sampled countrywide and from diverse socioeconomic backgrounds.

However, there are also limitations. First, the DID sample was limited to highly engaged users, especially for long-term evaluation. Thus, the results may not be generalizable to all users. Although baseline PA among all users was unknown, the DID sample likely included the most physically active users at baseline. Nevertheless, one-third of the DID sample (i.e., 89 baseline-level matched users) had a comparable daily step count to controls and the national average (29) at baseline, and their steps increased. Second, the smartphones were not always carried. Thus, step counts were likely underestimated (24). However, this systematic error was unlikely to differ between groups. Third, we only analyzed data from iPhone users. Because we did not have detailed information on the sample characteristics (e.g., income) of Android users, it is unknown how the exclusion of Android users altered the sample characteristics. The available functions in the app were exactly the same between iPhones and Android phones; thus, the effect should presumably be similar but needs to be examined in the future. Fourth, the app was based on a professional baseball league and may not be relevant to persons uninterested in baseball. Application of the model to other sports and entertainment (e.g., pop music fandom competing via billboard charts) merits further research and implementation. Finally, we did not investigate other potential benefits or harms. Although augmented reality (AR) games might entail risks, such as road traffic injuries, Pa-League Walk is not an AR or location-based. Thus, users do not need to view smartphones to accumulate steps. The average usage time is likely to be shorter than that of AR or location-based games.

CONCLUSIONS

Fandom-based gamification may promote PA among baseball fans at scale, including people with lower SES who are underrepresented in traditional health programs. Daily steps increased after app installation, and this was maintained for at least 9 months on average, longer than the previously tested gaming apps. This study highlights the importance of the integration strategy of behavioral design elements for long-term effect. Existing fandom or solidarity can be leveraged to promote PA among people with diverse interests and socioeconomic backgrounds.

We deeply appreciate the Pacific League Marketing Corporation for providing the data for independent research and Dr. Shiho Amagasa, Tokyo Medical University, for assisting with literature search. We are also immensely grateful to Dr. Yuhei Inoue, Manchester Metropolitan University, for comments that greatly improved the manuscript.

Contributors: M. K. conceptualized and designed the study, supervised all aspects of its implementation, performed the data analysis, and drafted the manuscript. As the corresponding author, M. K. has full access to all aspects of the research and writing process and takes final responsibility for the manuscript. H. H. and I. K. assisted with the conception and design of the study and drafting of the article. M. K., H. H., and I. K. gave input on the concept and design of the Pa-League Walk project, and solely the Pacific League Marketing Corporation managed the app. K. S. and M. T. assisted with writing code for analysis, interpreting data, and drafting the article. N. K. and I -M. L. supervised the analysis and interpretation of the data and helped draft and revise the manuscript. All authors provided revisions, contributed to the final manuscript, and agreed on its contents.

M. K. has received grants from the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Numbers 18H04071 and 19H03996) and the Meiji Yasuda Life Foundation of Health and Welfare (2018–2020). These funding bodies did not have any role in the study design, data collection, analysis, interpretation of the data, decision to publish, or preparation of the manuscript. For the remaining authors none were declared. The results of this study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

REFERENCES

1. 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 Health*. 2018;6(10):e1077–86.
2. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*. 2012;380(9838):219–29.
3. Patel MS, Benjamin EJ, Volpp KG, et al. Effect of a game-based intervention designed to enhance social incentives to increase physical activity among families: the BE FIT randomized clinical trial. *JAMA Intern Med*. 2017;177(11):1586–93.
4. Patel MS, Small DS, Harrison JD, 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;179(12):1624–32.
5. Edwards EA, Lumsden J, Rivas C, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open*. 2016;6(10):e012447.
6. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games*. 2013;1(1):e3.
7. Thaler RH, Sunstein CR. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven (CT): Yale University Press; 2008.
8. Patel MS, Asch DA, Rosin R, et al. Framing financial incentives to increase physical activity among overweight and obese adults: a randomized, controlled trial. *Ann Intern Med*. 2016;164(6):385–94.
9. Münscher R, Vetter M, Scheuerle T. A review and taxonomy of choice architecture techniques. *J Behav Dec Making*. 2016;29(5):511–24.
10. Forberger S, Reisch L, Kampfmann T, Zeeb H. Nudging to move: a scoping review of the use of choice architecture interventions to promote physical activity in the general population. *Int J Behav Nutr Phys Act*. 2019;16:77.
11. Romeo A, Edney S, Plotnikoff R, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. *J Med Internet Res*. 2019;21(3):e12053.
12. Sardi L, Idri A, Fernandez-Aleman JL. A systematic review of gamification in e-Health. *J Biomed Inform*. 2017;71:31–48.
13. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. *JMIR Mhealth Uhealth*. 2018;6(3):e53.
14. Laranjo L, Ding D, Heleno B, et al. Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression. *Br J Sports Med*. 2021;55:422–32.
15. Wann DL. Understanding the positive social psychological benefits of sport team identification: the Team Identification—Social Psychological Health Model. *Group Dyn Theory Res Pract*. 2006;10(4):272–96.
16. Inoue Y, Berg BK, Chelladurai P. Spectator sport and population health: a scoping study. *J Sport Manag*. 2015;29(6):705–25.
17. Hunt K, Wyke S, Gray CM, et al. A gender-sensitised weight loss and healthy living programme for overweight and obese men delivered by Scottish Premier League football clubs (FFIT): a pragmatic randomised controlled trial. *Lancet*. 2014;383(9924):1211–21.
18. Gray CM, Wyke S, Zhang R, et al. Long-term weight loss trajectories following participation in a randomised controlled trial of a weight management programme for men delivered through professional football clubs: a longitudinal cohort study and economic evaluation. *Int J Behav Nutr Phys Act*. 2018;15:60.
19. Macromill, Mitsubishi UFJ Research and Consulting. Sports Marketing Basic Survey 2019. 2019 [cited 2020 Jul 30]. Available from: https://www.macromill.com/wp-content/uploads/files/press/release/pdf/20191011_macromill.pdf.
20. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med*. 2013;46(1):81–95.
21. Yates T, Haffner SM, Schulte PJ, et al. Association between change in daily ambulatory activity and cardiovascular events in people with impaired glucose tolerance (NAVIGATOR trial): a cohort analysis. *Lancet*. 2014;383(9922):1059–66.
22. Saint-Maurice PF, Troiano RP, Bassett DR Jr, et al. Association of daily step count and step intensity with mortality among US adults. *JAMA*. 2020;323(12):1151–60.
23. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA*. 2015;313(6):625–6.
24. Amagasa S, Kamada M, Sasai H, et al. How well iPhones measure steps in free-living conditions: cross-sectional validation study. *JMIR Mhealth Uhealth*. 2019;7(1):e10418.
25. Oka K. Reliability and validity of the stages of change for exercise behavior scale among middle-age adults. *Jpn J Health Promot*. 2003;5(1):15–22.
26. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA*. 2014;312(22):2401–2.
27. Daw JR, Hatfield LA. Matching and regression to the mean in difference-in-differences analysis. *Health Serv Res*. 2018;53(6):4138–56.
28. Ministry of Health, Labour, and Welfare of Japan. The National Health and Nutrition Survey in Japan, 2013. 2015 [cited 2020 Jul 30]. Available from: https://www.mhlw.go.jp/bunya/kenkou/eiyou/h25-houkoku.html.
29. Ministry of Health, Labour, and Welfare of Japan. The National Health and Nutrition Survey in Japan, 2016. 2017 [cited 2021 Jun 30]. Available from: https://www.mhlw.go.jp/bunya/kenkou/eiyou/h28-houkoku.html.
30. Tudor-Locke C, Craig CL, Thyfault JP, Spence JC. A step-defined sedentary lifestyle index: <5000 steps/day. *Appl Physiol Nutr Metab*. 2013;38(2):100–14.
31. Howe KB, Suharlim C, Ueda P, Howe D, Kawachi I, Rimm EB. Gotta catch'em all! Pokémon GO and physical activity among young adults: difference in differences study. *BMJ*. 2016;355:i6270.
32. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. 2000;55(1):68–78.
33. Wyke S, Bunn C, Andersen E, et al. The effect of a programme to improve men's sedentary time and physical activity: The European Fans in Training (EuroFIT) randomised controlled trial. *PLoS Med*. 2019;16(2):e1002736.
34. Pringle A, Zwolinsky S, McKenna J, Daly-Smith A, Robertson S, White A. Effect of a national programme of men's health delivered in English Premier League football clubs. *Public Health*. 2013;127(1):18–26.
35. Nielsen Sports. World Football Report 2018. 2018 [cited 2021 Mar 8]. Available from: https://www.nielsen.com/us/en/insights/report/2018/world-football-report/.
36. Garber CE, Allsworth JE, Marcus BH, Hesser J, Lapane KL. Correlates of the stages of change for physical activity in a population survey. *Am J Public Health*. 2008;98(5):897–904.
37. Maslow AH. A theory of human motivation. *Psychol Rev*. 1943;50:370–96.
38. World Health Organization (WHO). *Global Action Plan on Physical Activity 2018–2030: More Active People for a Healthier World*. Geneva (Switzerland): WHO; 2018.