Personalizing Mobile Fitness Apps using Reinforcement Learning

Mo Zhou
Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
mzhou@berkeley.edu

Yonatan Mintz
Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
ymintz@berkeley.edu

Yoshimi Fukuoka
Department of Physiological Nursing
Institute for Health & Aging, School of Nursing
University of California, San Francisco, CA, USA
yoshimi.fukuoka@ucsf.edu

Ken Goldberg
Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
goldberg@berkeley.edu

Elena Flowers
Department of Physiological Nursing, School of Nursing
University of California, San Francisco, CA, USA
elena.flowers@ucsf.edu

Philip Kaminsky
Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
kaminsky@berkeley.edu

Alejandro Castillejo
Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
castillejo.alejandro@berkeley.edu

Anil Aswani
Department of Industrial Engineering and Operations Research
University of California, Berkeley, CA, USA
aaswani@berkeley.edu

ABSTRACT
Despite the vast number of mobile fitness applications (apps) and their potential advantages in promoting physical activity, many existing apps lack behavior-change features and are not able to maintain behavior change motivation. This paper describes a novel fitness app called CalFit, which implements important behavior-change features like dynamic goal setting and self-monitoring. CalFit uses a reinforcement learning algorithm to generate personalized daily step goals that are challenging but attainable. We conducted the Mobile Student Activity Reinforcement (mSTAR) study with 13 college students to evaluate the efficacy of the CalFit app . The control group (receiving goals of 10,000 steps/day) had a decrease in daily step count of 1,520 (SD $\pm$ 740) between baseline and 10-weeks, compared to an increase of 700 (SD $\pm$ 830) in the intervention group (receiving personalized step goals). The difference in daily steps between the two groups was 2,220, with a statistically significant $p = 0.039$.

HUMANIZE '18, March 11, 2018, Tokyo, Japan

ACM Classification Keywords
H.5.2. User Interfaces: User-centered design; I.2.6 Artificial Intelligence: Learning; K.4.1 Computers and Society: Public Policy Issues: Computer-related healthcare issues; J.4 Social and Behavioral Sciences: Psychology

Author Keywords
Physical activity; interface design; mobile app; fitness app; goal setting; personalization.

INTRODUCTION
Regular physical activity (e.g., walking or running) is an important factor in preventing the development of chronic diseases like type 2 diabetes, cardiovascular disease, depression, and certain types of cancer [33, 55, 56]. Because of its importance in maintaining good health, the 2008 Physical Activity Guidelines for Americans recommend that adults engage in at least 150 minutes a week of moderate-intensity physical

Figure 1. Screenshots of the main tabs of the CalFit app are shown, including the (a) splash screen, (b) home tab, (c) history tab, and (d) contact tab.

activity or 75 minutes a week of vigorous-intensity aerobic physical activity [51, 55]. However, about 50% of adults in the U.S. [15] are physically inactive. In fact, over 3 million deaths worldwide are attributed to physical inactivity [54].

Given the high prevalence of physical inactivity, it is necessary to develop new cost-effective, scalable approaches to increase physical activity. One promising direction is the use of smartphones in the delivery and personalization of programs that motivate individuals to increase their physical activity. Over 40% of adults worldwide and 77% of adults in the U.S. own a smartphone [45]. Smartphones have powerful computation and communication capabilities that enable the use of machine learning and other data-driven analytics algorithms for personalizing the physical activity programs to each individual. Furthermore, the past several generations of smartphones integrate reliable activity tracking features [2, 14, 18, 25], which makes possible the real-time collection of fine-grained physical activity data from each individual.

Though many smartphone applications (apps) for fitness have been developed, systematic reviews [8, 10, 39, 53] of mobile fitness apps found an overall lack of persuasive attributes that are needed for the general public to maintain exercise motivation through continued use of the app. These reviews [10, 53] also identified a lack of experimental validation for the efficacy of specific features implemented in mobile fitness apps. For instance, recent studies [28, 36, 47] have shown that constant step goals provided by existing apps and devices are ineffective in increasing physical activity and such a one-size-fits-all approach could even be harmful for some people. Therefore, maintaining user participation and motivation is a core challenge in developing effective physical activity intervention platforms, and the personalization of goals within fitness apps through intelligent user interfaces [16, 21, 30, 48, 52] has shown promise in promoting healthy behavior. Simple heuristics , such as setting the future goal to be the 60th percentile of the steps taken in the past 10 days, has shown to be effective in promoting physical activity [1]. But few studies have investigated the potential of using more complicated Machine Learning-based approaches to set personalized step goals.

In this paper, we introduce a novel fitness app on the iOS platform, CalFit, which automatically sets personalized, adaptive daily step goals and adopts behavior-change features such as self-monitoring. The daily step goals are computed using a reinforcement learning algorithm [5, 40] adapted to the context of physical activity interventions: Our app uses inverse reinforcement learning to construct a predictive quantitative model for each user, and then uses this estimated model in conjunction with reinforcement learning to generate challenging but realistic step goals in an adaptive fashion. We conducted a pilot study with 13 college students to demonstrate the efficacy of our app and the personalized adaptive step goal algorithm in promoting physical activity.

We first discuss related work and the theory of goal setting in relation to behavior change. Next we describe the designed elements. Our contributions toward the app design include translating elements and features from the theory of goal setting into interface design choices for mobile fitness apps, as well as the design of a reinforcement learning algorithm that generates personalized step goals for users. Next we describe our contributions towards experimental validation of the efficacy of our app design, through conducting the Mobile Student Activity Reinforcement (mSTAR) study.

RELATED WORK

In this section, we review work on the intersection of mobile technologies and behavior modification programs. First, we describe key studies showing the efficacy of combining mobile technologies with clinical coaching to increase physical activity. Next, we describe behavior change features and their use in the design of mobile fitness apps. Finally, we survey the theory of goal-setting. Identified weaknesses in existing apps and ideas on the theory of behavior change are used to inform our design of the CalFit app.

Smartphone-based Clinical Trials

Physical activity interventions that involve multiple in-person coaching sessions are costly and labor-intensive, and so researchers have evaluated the feasibility and efficacy of lower-cost interventions where the number of coaching sessions are reduced (but not eliminated) in parallel with the introduction of mobile technologies (e.g., smartphone apps, digital pedometers, activity trackers) [7, 11, 12, 13, 17, 19, 20, 23, 26, 29, 32, 44, 46]. These studies ranged in size from about 10 to several hundred participants. Both smartphones and personal digital assistants (PDA's) were used to deliver these interventions, and the interface outputs were predominately text with some interventions involving simple graphic comparisons to goals.

These interventions featured different levels of interactivity, ranging from general weekly text messages to customized text messages based on real-time monitoring of physical activity and other additional inputs. For instance, the mobile weight loss program in [11] used weekly input from overweight children to send computer-generated text messages. Most studies [7, 11, 12, 13, 17, 20, 23, 29, 32, 44] asked participants to self-report dietary, weight, and exercise data. A smaller number of studies [19, 26] have explored the use of automated collection of exercise data either through accelerometer data that is wirelessly transmitted via Bluetooth to a smartphone [26] or the use of digital pedometers to collect step data [19].

Most of these studies had outcomes of a statistically significant decrease in weight or a statistically significant increase in physical activity [7, 11, 13, 17, 20, 23, 29, 32, 44], supporting the potential advantages of mobile-based physical activity interventions. However, none of these studies relied solely on mobile technology. All of these studies involved in-person coaching sessions during the intervention (though the number of coaching sessions was lower than in traditional behavior modification programs) and either used objectively measured outcomes using an additional device or self-reported outcomes. The weight or exercise goals in these interventions were manually set by the participants or the clinicians.

Mobile Fitness Apps and Behavior Change Features

Mobile fitness apps have the potential to be a scalable way of disseminating behavior change interventions in a cost-effective manner. In addition to being able to deliver interventions through wireless internet and messaging connectivity, smartphones can also leverage in-built tools like GPS, digital accelerometers, and cameras to objectively measure (as opposed to self-reported data) health parameters. However, systematic reviews [8, 10, 39, 53] of current mobile fitness apps found a lack of features that can effectively initiate and maintain the behavioral changes necessary to increase physical activity.

The low efficacy of current mobile fitness apps is due primarily to this lack of inclusion of important features based on behavioral theory [8, 10, 39, 53]. Examples of key behavior change features include: objective outcome measurements, self-monitoring, personalized feedback, behavioral goal-setting, individualized program, and social support. In particular, researchers recommend that self-monitoring should be conducted regularly and in real-time, so as to target activity with precise tracking information and emphasize performance successes. In addition, personalized feedback is most effective when it is specific, such as in comparing current performance to past accomplishments and previous goals.

Goal Setting

Goal setting is a critical factor for facilitating behavior change [9, 37]. Prior studies using persuasive technology usually assigned a fixed goal to all participants (e.g., 10,000 steps per day) [3, 28, 36], but a fixed goal fails to capture the differences between participants (different baseline physical activity level, reaction to goals, etc). Conversely, personalized goal setting have the potential to increase the effectiveness of physical activity interventions. Simple heuristics, such as setting the future goal to be the 60th percentile of the steps taken in the past 10 days, has shown to be effective in promoting physical activity [1]. But few studies have investigated the potential of using more complicated Machine Learning-based approaches to set personalized step goals.

In recent years, human-computer interaction (HCI) studies have investigated interface design for goal-setting. Munson et al. [41] developed a smartphone app that implements primary (base) and secondary (stretch) weekly goals and found that such a personalized goal-setting approach can be beneficial. However, the app lacks an explicit algorithm to help participants set "sweet spot" goals based on their past behavior. DStress [34] algorithmically sets daily goals based on previous performance, where if the daily goal is achieved for the day, then a more difficult goal is assigned for the following day and vice versa. Though this can effectively set adaptive goals, the goals for high variance targets (like steps) can be highly variate, which leads to reduced intervention impact. For example, if a participant normally walks 8,000 steps but walks 1,000 steps on one day, then using the 1,000 value as the baseline to set the step goal for the following day will lead to a too-easy goal. A more comprehensive algorithm is needed to incorporate all previous performance information to decide the "sweet spot" of future goals in a personalized fashion. In this paper, we describe a novel algorithm based on Reinforcement Learning that set goals 'smartly' by first learning the behaviors of each participant and then determines the most effective future goal in an adaptive fashion.

THE CALFIT APP

CalFit is a mobile fitness app that uses key behavior change features to improve effectiveness. It combines a personalized goal setting algorithm and a structured interface with regular self-monitoring and feedback to provide an adaptive

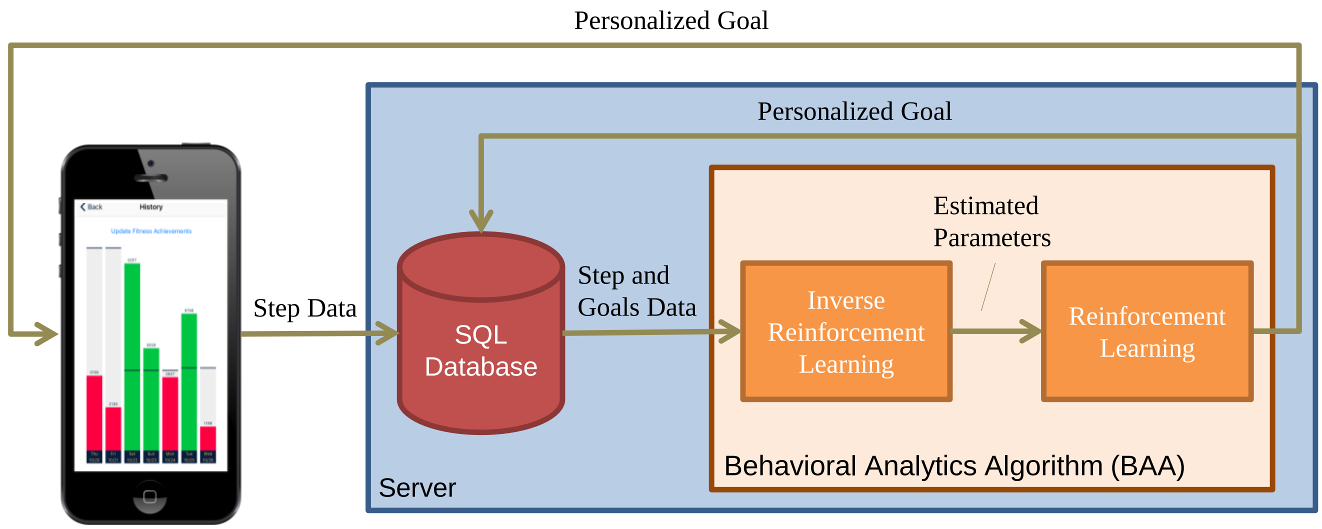

Figure 2. The CalFit app interface uploads step data to a SQL database on a server, and the stored step and goals data is accessed by the Behavioral Analytics Algorithm (BAA) comprised of inverse reinforcement learning to estimate model parameters describing the user and followed by reinforcement learning to compute personalized step goals that will maximize the user's future physical activity. The personalized step goals are stored in the SQL database and communicated to the user via the CalFit app interface.

and individualized physical activity intervention. This section discusses the design of the interface, communication, and computation elements of our app, which are shown in Figure 2.

Interface

The CalFit app interface is built for the iOS platform. Upon opening the app interface, the user first sees the splash screen (Figure 1a) and then lands on the home tab (Figure 1b). On the home tab, the user can find his/her step goal for the day and the steps done so far today. The steps are tracked in real-time using the built-in health chip on the iPhone and are updated every 10-minutes. (Accuracy of step data collected by the built-in health chip on the iPhone and other smartphones has been validated by several studies [2, 14, 18, 25].) This design facilitates direct comparison between daily step goals and objectively measured daily steps in order to enhance self-monitoring.

There are two icons at the bottom of the home tab. If the left icon on the home tab is clicked, the user is shown the history tab (Figure 1c) that displays a barplot outlining the user's performance in the past 7 days. The black lines on each bar represent the step goal, and the height of each bar represents the actual measured steps. If the user achieved the goal, then the bar is green. If the user did not achieve the goal, then the bar is red. This tab is designed to provide a quick, yet comprehensive, visualization of the user's past performance, allowing the user to quickly identify days of successes and failures. If the right icon on the home tab is clicked, the user is directed to the contact tab (Figure 1d), where they can type in a message and send it to the research team regarding their concerns, app bugs, etc.

Behavioral Analytics Algorithm (BAA)

Automated goal setting is a crucial component of the CalFit app. To set personalized goals that are challenging yet attainable for each user, we use a reinforcement learning algorithm [5, 40] that we have adapted to the context of physical activity interventions. The Behavioral Analytics Algorithm (BAA) uses inverse reinforcement learning to construct a predictive quantitative model for each participant, based on the historical step and goal data for that user; then, it uses the estimated model with reinforcement learning to generate challenging but realistic step goals in an adaptive fashion.

Below, we elaborate upon the mathematical formulations underlying these steps of BAA. Since the BAA algorithm does calculations for each user independently of the calculations for other users, our description of the algorithm (and accompanying models) is focused on calculations for a single user.

Stage 0 – Predictive Model of User's Step Activity

Our predictive model is based on a model from [5, 40] for predicting weight loss based on steps and diet, and we have adapted that model to the specific case of only predicting step activity. Let the subscript t denote the value of a variable on the t-th day of using the app, and define the function $(x)^-$ as

$$(x)^- = \begin{cases} x, & \text{if } x \le 0 \\ 0, & \text{if } x > 0 \end{cases} \tag{1}$$

Our predictive model for the number of steps that the user takes on the t-th day is

$$u_t = \arg\max_{u \ge 0} \; -(u-u_b)^2 + p_t \cdot (u-g_t)^-, \tag{2}$$

where u_t is the number of steps the user (subconsciously) decides to take, $u_b \in \mathbb{R}_+$ is a parameter describing the user's natural (or baseline) level of steps in a day, and $p_t \in \mathbb{R}_+$ is a parameter that quantitatively characterizes the user's responsiveness to the goal $g_t \in \mathbb{R}_+$.

The general idea of (2) is that users make decisions to maximize their utility or happiness related to several objectives. The $-(u-u_b)^2$ term means a user has an ideal level of steps they prefer to take in a day, wherein the user is implicitly trading off a small number of steps in a day (and the dissatisfaction accompanied by physical inactivity) with a large number of steps in a day (and the effort and time required to achieve many steps). The parameter u_b quantifies this baseline number of steps that achieves this tradeoff for the user. The $p_t \cdot (u-g_t)^-$ term means a user gets increasing happiness the closer their steps are to the goal g_t, and p_t describes the rate of increase in happiness as the steps get closer to the goal; however, this model says that exceeding the goal results in no additional happiness. A more complex model would include a term to describe an increase in happiness as the goal is exceeded, but a detailed study [5] found that not including this additional term still produced a model with high prediction accuracy.

There is one additional component to our predictive model. Equation (2) describes how a user decides the number of steps to take on the t-th day. The theory of goal setting [9, 37] recognizes that the effectiveness of goals can increase or decrease over time, depending on the level of the goals and whether or not an individual was able to meet the goals. To quantify these effects, our predictive model includes

$$p_{t+1} = \gamma \cdot p_t + \mu \cdot \mathbf{1}\left(u_t \geq g_t\right), \tag{3}$$

where $\gamma \in (0,1)$ characterizes the user's learned helplessness, $\mu \in \mathbb{R}_+$ quantifies the user's self-efficacy, and $\mathbf{1}(\cdot)$ is the indicator function. Self-efficacy is defined as a user's beliefs in their capabilities to successfully execute courses of action, and it plays an essential role in the theory of goal setting [9, 37]. Self-efficacy influences a variety of health behaviors, including physical activity [31, 38]. Though γ will be different for each individual, the past study [5] found that setting $\gamma = 0.85$ generated models with high prediction accuracy.

There are several points of intuition about (3). The term $\mu \cdot \mathbf{1}\left(u_t \geq g_t\right)$ describes the relationship between self-efficacy and meeting goals. When a user achieves a goal, $\mathbf{1}\left(u_t \geq g_t\right)$ is one and p_{t+1} increases by μ. Achieving a goal increases the user's self-efficacy, leading to increased steps on future days. But if the user misses a goal, then $\mathbf{1}\left(u_t \geq g_t\right)$ is zero and p_{t+1} does not increase. Not achieving a goal decreases the user's self-efficacy, leading to lower steps in the future. The term $\gamma \cdot p_t$ describes the phenomenon whereby learned helplessness reduces the utility or happiness an individual achieves for achieving goals. Consequently, (3) captures the interplay between increasing self-efficacy from meeting specific goals with the decrease in self-efficacy from learned helplessness.

Stage 1 – Inverse Reinforcement Learning

The BAA algorithm first uses inverse reinforcement learning to estimate the parameters u_b, p_t, μ in the predictive model (2), (3) for a user. Denoting n measurements of the user's step counts at times t_i as $\tilde{u}_{t_i}$, for $i = 1, \ldots, n$, our measurement model $\tilde{u}_{t_i} = u_{t_i} + \varepsilon_i$ is that the observed step counts $\tilde{u}_{t_i}$ deviate from the step counts chosen in the predictive model u_{t_i} by an additive zero mean random variable ε_i. The study [5] found that assuming ε_i has a Laplacian distribution led to an easily computable formulation and generated accurate predictions.

Under the above setup, the inverse reinforcement learning problem [6, 24, 35, 42] is equivalent to estimating the model parameters u_b, p_t, μ. This problem can be formulated as a log-likelihood maximization [5, 40]. If we define H to be the duration of the intervention, then we can write this estimation problem as a bilevel optimization problem

$$\begin{aligned}
\min\ & \sum_{i=1}^{n} \left|u_{t_i} - \tilde{u}_{t_i}\right| \\
\text{s.t. }& u_t = \arg\max_{u \geq 0}\ -(u-u_b)^2 + p_t \cdot (u-g_t)^- \\
& p_{t+1} = \gamma \cdot p_t + \mu \cdot \mathbf{1}\left(u_t \geq g_t\right) \\
& 0 \leq p_t, \mu \leq \mathsf{UB}_p \\
& 0 \leq u_t, u_b \leq \mathsf{UB}_u
\end{aligned} \tag{4}$$

where the constraints hold for $t = 1, \ldots, H$, and $\mathsf{UB}_p, \mathsf{UB}_u$ are constants that are upper bounds on the possible values. Existing numerical optimization software is not able to solve the above problem, but we can rewrite it as a mixed-integer linear program (MILP) [5, 40]. Let δ be a small positive constant, and M be a large positive constant. The above optimization problem can be rewritten as the following MILP:

$$\begin{aligned}
\min\ & \sum_{i=1}^{n_u} a_{t_i} \\
\text{s.t. }& -a_{t_i} \leq u_{t_i} - \tilde{u}_{t_i} \leq a_{t_i} \\
& u_t = \tfrac{1}{2}(\lambda_{1,t} + \lambda_{3,t}) + u_b \\
& 0 \leq \lambda_{3,t} \leq p_t \\
& (g_t - \delta) - Mx_{1,t} \leq u_t \leq g_t - \delta + M(1 - x_{1,t}) \\
& (g_t - \delta) - M(1 - x_{2,t}) \leq u_t \leq g_t + \delta + M(1 - x_{2,t}) \\
& (g_t + \delta) - M(1 - x_{3,t}) \leq u_t \leq g_t + \delta + Mx_{3,t} \\
& p_t - M \cdot (1 - x_{t,1}) \leq \lambda_{3,t} \leq M \cdot (1 - x_{3,t}) \\
& u_t \leq My_{u,1} \\
& \lambda_{1,t} \leq M \cdot (1 - y_{u,t}) \\
& p_{t+1} \geq \gamma \cdot p_t \\
& p_{t+1} \leq \gamma \cdot p_t + M \cdot (1 - x_{t,1}) \\
& p_{t+1} \geq \gamma \cdot p_t + \mu - M \cdot x_{t,1} \\
& p_{t+1} \leq \gamma \cdot p_t + \mu \\
& x_{t+1,1} \geq x_{t,1} - \mathbf{1}(g_{t+1} - g_t < 0) \\
& x_{t+1,2} \leq x_{t,2} + \mathbf{1}(g_{t+1} - g_t < 0) \\
& x_{t+1,3} \leq x_{t,3} + \mathbf{1}(g_{t+1} - g_t < 0) \\
& x_{t,1} + x_{t,2} + x_{t,3} = 1 \\
& y_{u,t}, x_{t,1}, x_{t,2}, x_{t,3} \in \{0, 1\} \\
& \lambda_{1,t} \geq 0 \\
& 0 \leq p_t, \mu \leq \mathsf{UB}_p \\
& 0 \leq u_t, u_b \leq \mathsf{UB}_u
\end{aligned} \tag{5}$$

where the constraints hold for $t = 1, \ldots, H$ and $i = 1, \ldots, n$. The above MILP can be easily solved using standard optimization software [4, 22, 27].

Stage 2 – Reinforcement Learning

Under our setup, the reinforcement learning problem [40, 49, 50] for computing an optimal set of personalized goals for the user is equivalent to performing a direct policy search using the estimated model parameters $\hat{u}_b, \hat{p}_0, \hat{\mu}$ computed by solving (5). Adapting the solution in [40] to the current context of choosing an optimal sequence of step goals leads to a MILP:

$$
\begin{aligned}
\max\ & u_{min} \\
\text{s.t. } & u_{min} \leq u_t, \text{ for } t > T \\
& -\delta \leq u_t - \hat{u}_t \leq \delta, \text{ for } t \leq T \\
& -\delta \leq p_t - \hat{p}_t \leq \delta, \text{ for } t \leq T \\
& u_t = \tfrac{1}{2}(\lambda_{1,t} + \lambda_{3,t}) + \hat{u}_b \\
& 0 \leq \lambda_{3,t} \leq p_t \\
& (g_t - \delta) - Mx_{1,t} \leq u_t \leq g_t - \delta + M(1 - x_{1,t}) \\
& (g_t - \delta) - M(1 - x_{2,t}) \leq u_t \leq g_t + \delta + M(1 - x_{2,t}) \\
& (g_t + \delta) - M(1 - x_{3,t}) \leq u_t \leq g_t + \delta + Mx_{3,t} \\
& p_t - M(1 - x_{t,1}) \leq \lambda_{3,t} \leq M(1 - x_{3,t}) \\
& u_t \leq My_{u,1} \\
& \lambda_{1,t} \leq M(1 - y_{u,t}) \\
& p_{t+1} = \gamma p_t + \hat{\mu}(1 - x_{1,t}), \text{ for } t > T \\
& x_{t+1,1} \geq x_{t,1} - g_{ind,t}, \text{ for } t > T \\
& x_{t+1,2} \leq x_{t,2} + g_{ind,t}, \text{ for } t > T \\
& x_{t+1,3} \leq x_{t,3} + g_{ind,t}, \text{ for } t > T \\
& x_{t,1} + x_{t,2} + x_{t,3} = 1 \\
& g_{t+1} - g_t \leq M(1 - g_{ind,t}), \text{ for } t > T \\
& g_{t+1} - g_t \geq -Mg_{ind,t}, \text{ for } t > T \\
& y_{u,t}, x_{t,1}, x_{t,2}, x_{t,3}, g_{ind,t} \in \{0,1\}, \text{ for } t > T \\
& \lambda_{1,t} \geq 0 \\
& 0 \leq p_t \leq \mathsf{UB}_p \\
& 0 \leq u_t \leq \mathsf{UB}_u
\end{aligned}
\tag{6}
$$

where T is the current time, and the remaining constraints hold for $t = 1, \ldots, H$ and $i = 1, \ldots, n$. The intuition is that the above MILP picks future goals in order to maximize the smallest number of steps on any given day in the future, and the reason for this choice is that in our simulations we found that this objective function choice led to the largest increases (as compared to other possible objective function choices) in physical activity. Moreover, the above MILP can be easily solved using standard optimization software [4, 22, 27].

Feedback via Push Notification

Using the BAA algorithm, the CalFit app is able to adaptively set personalized step goals for users. To optimize the impact of this goal-setting algorithm, we implemented feedback features via iOS push notifications. Each user receives at most two push notifications each day. The first push notification is received by every user at 8:00am, and it notifies the users about their goal for the day. The second push notification at 8:00pm is only received by users who successfully achieved their step goal for the day. Note the standard iOS push notification is used (i.e., appears in both the landing page and the recent notifications tab), and a user receives push notifications regardless of whether or not the CalFit app interface is on; and if the push notification is clicked, it will lead to the homepage of the app interface. The benefit of sending push notifications is two-fold: First of all, we want to constantly engage the users to implicitly remind them to continue using the app interface. This is particularly important for fully automated physical activity interventions since users have a lower intention to adhere due to the lack of in-person coaching sessions. Secondly, the congratulating push notifications can be seen as customized assessment/feedback to users on their daily performance.

Implementation Details

The CalFit app consists of two parts: The interface of the iOS app (including push notifications) and the BAA dynamic goal setting algorithm. The backend of the CalFit app was implemented via the Parse API [43] running on an Intel Xeon E5-2650 v3 2.3GHz Turbo server with 16GB RAM. The server was running CentOS 6.6, and the data was stored in an SQLite database on the same server. The BAA algorithm was written in Python, and the MILP's were solved using Gurobi [22]. The running time for BAA to recommend goals for a single user was less than one second on average, which is in line with the benchmarks from [5] for personalizing a weight intervention.

THE mSTAR STUDY

To experimentally evaluate the efficacy of the CalFit app and personalized goal setting using the BAA algorithm, we conducted the Mobile Student Activity Reinforcement (mSTAR) study with college students in University of California, Berkeley (UCB). The main research question was: Does setting personalized step goals increase user's steps compared to fixed step goals? The secondary research question was: Does setting personalized step goals improve adherence? The study was approved by the Committee for Protection of Human Subjects of the University of California, Berkeley (IRB Number 2016-03-8609) in July 2016. All participants provided written informed consent prior to study enrollment.

Methodology

To evaluate the above hypotheses, we designed the app so that each user is randomly assigned to either the control group or the intervention group upon joining the study. Users in the control group received constant step goals of 10,000 steps everyday during the trial, whereas users in the intervention group received personalized step goals computed by the BAA algorithm. Both groups received the morning and evening push notifications.

Participants

We recruited UCB students by sending email announcements to departments. Recruitment started in January 2017 and ended in February 2017. Interested students were directed to complete an online survey to assess eligibility, and eligible students were encouraged to sign-up for an in-person session to complete enrollment in the study and install the app. The students were randomly assigned to either the control group or the intervention group upon installation of the app.

The inclusion criteria was: being a full-time UCB student, intent to become physically active, own an iPhone 5s or newer

	All Users (N=13)	Control (N=7)	Intervention (N=6)	p-value
	Mean (± SD)	**Mean (± SD)**	**Mean (± SD)**	
Baseline daily average steps	6,163 ± 1,822	6,829 ± 2,023	5,387 ± 1,309	0.16
Age (years)	22.2 ± 2.9	21.6 ± 2.3	23.0 ± 3.5	0.40
Weight (kg)	70.4 ± 23.9	73.7 ± 31.9	66.5 ± 20.8	0.61
	% (N)	**% (N)**	**% (N)**	
Gender				0.88
Male	23.1 (3)	14.3 (1)	33.3 (2)	
Female	76.9 (10)	85.7 (6)	66.6 (4)	
Ethnicity				0.85
Asian	23.1 (3)	28.6 (2)	16.7 (1)	
Hispanic/Latino	15.4 (2)	14.3 (1)	16.7 (1)	
White (non-Hispanic)	23.3 (3)	28.6 (2)	16.7 (1)	
Other	38.5 (5)	28.6 (2)	50.0 (3)	
Marital Status				1.00
Currently Married/Cohabitating	7.7 (1)	14.3 (1)	0.0 (0)	
Never Married	92.3 (12)	85.7 (6)	100.0 (6)	
Divorced/Widowed	0.0 (0)	0.0 (0)	0.0 (0)	
Year in School				0.52
Freshman	0.0 (0)	0.0 (0)	0.0 (0)	
Sophomore	15.4 (2)	28.6 (2)	0.0 (0)	
Junior	30.8 (4)	28.6 (2)	33.3 (2)	
Senior	23.1 (3)	14.3 (1)	33.3 (2)	
Graduate	30.8 (4)	28.6 (2)	33.3 (2)	
Own a Dog				1.00
Yes	7.7 (1)	14.3 (1)	0.0 (0)	
No	92.3 (12)	85.7 (6)	100.0 (6)	
Transportation to Work				0.43
Car	23.1 (3)	28.6 (2)	16.7 (1)	
Public Transportation	7.7 (1)	14.3 (1)	0.0 (0)	
Walk	61.5 (8)	42.9 (3)	83.3 (5)	
Bicycle	7.7 (1)	14.3 (1)	0.0 (0)	

Table 1. Comparison of Baseline Characteristics shows that the differences between participants in the control and intervention groups were not statistically significant, which is expected since participants were randomly assigned to groups.

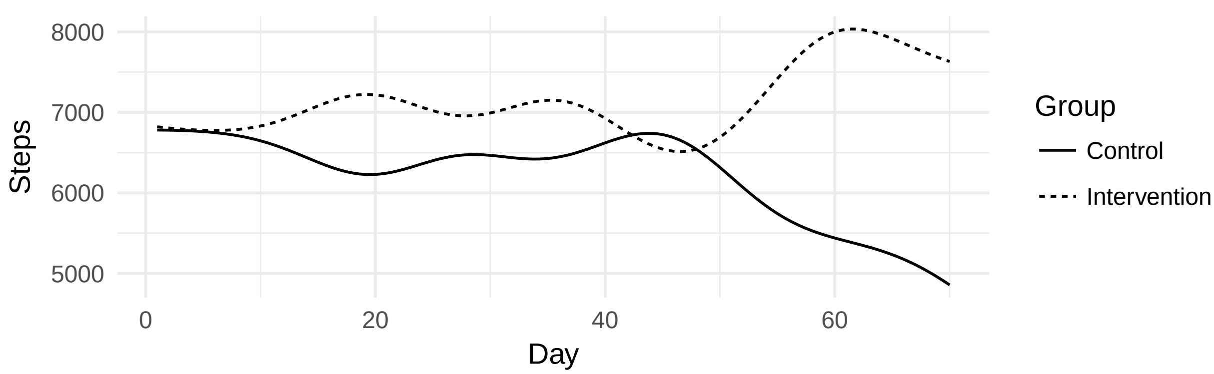

Figure 3. The objectively measured daily steps of the control group and the intervention group over the 10-week study period show the statistically significant difference in the number of daily steps at the end of the study. The plotted values are computed by averaging the raw data over each user in the corresponding group, adjusting the baseline value based on the value computed from the LMM model, and then smoothing the data using a standard (nonparametric) Nadaraya-Watson estimator.

model, and willing to carry the iPhone during the study period. The exclusion criteria was: preexisting health conditions that may make participation unsafe, having participated recently in a physical activity or weight loss intervention, and regularly taking 20,000 steps in a day. We excluded students who took 20,000 steps per day because it is not possible to increase activity by using our app if they were at that activity level (since the BAA algorithm uses 20,000 steps as the upper bound for the goal), and the procedure was that students satisfying the other criteria were enrolled and then excluded if 20,000 steps was observed in the step data collected.

Study Procedure

Eligible users were required to attend two 15-minute in-person sessions (one at baseline and one at study conclusion). The first in-person session occurred in January 2017 and the second occurred in May 2017. During the first in-person session, a trained research staff member installed the CalFit app on users' phones and advised them to carry the phone on their person everyday during their participation in the study. The users were randomly assigned to either the control group or the intervention group upon app installation. No other in-person sessions were conducted during the study period to simulate a fully smartphone application-based study environment, which is similar to the environment of most fitness apps.

The users started a 1-week run-in period after the first in-person session. All users received identical daily step goals of: 3000, 3500, 4000, 4500, 5000, 5500, and 6000 steps. This set of adaptive run-in steps goals was designed to engage the users in using the application regularly. Also, the morning and the evening push notifications were sent to all eligible users. Because the same step goals were provided to both the control and intervention groups, we were able to collect run-in daily steps data when both groups received identical treatment.

After the 1-week run-in period, the daily step goals for users in the control group (N=7) were set to 10,000 steps/day through the CalFit app, whereas the daily step goals for users in the intervention group (N=6) were set by the BAA algorithm. The BAA algorithm was applied every week (to mitigate the impact of large step variance), and it computes the step goals for the following 7 days. Both groups received morning and evening push notifications. The study lasted for 10-weeks, and participants could earn up to a $25 Amazon gift card for completing all parts of the study, including attending a final in-person session.

RESULTS

Table 1 shows the baseline characteristics of the participants. The overall mean age was 22.2 (SD $\pm$ 2.9) years and 77% of the participants were female. The baseline mean daily step in the control group was slightly higher than that in the intervention group, but the difference is not statistically significant (6,829 steps versus 5,387 steps, respectively; $p = 0.16$). The p-values in Table 1 were computed using t-tests for continuous variables and χ^2-tests for categorical variables.

Physical Activity Outcomes

The primary outcome of the study is the objectively measured daily steps from baseline to 10-weeks. We conducted our statistical analysis of the primary outcome of daily steps using a linear mixed-effects model (LMM) [31, 33, 38] with random effects for each individual of random slope and random intercept, and fixed effects of time, intervention group, and interaction term of time and intervention group. This analysis found that the control group had a decrease in daily step count of 1520 (SD $\pm$ 740) steps between baseline and 10-weeks, compared to an increase of 700 (SD $\pm$ 830) steps in the intervention group. The difference in daily steps between the two groups was 2220 ($p = 0.039$) with a 95% confidence interval of (100, 4480), which is a statistically significant difference. The step goals computed by the BAA algorithm were on average between 6,000 steps and 8,000 steps. They varied between different users and days resulting from its adaptive and personalized nature.

Figure 3 shows the change in daily steps over the 10-week study period, and for fair comparison we baseline-adjusted the plotted steps by adding the coefficient corresponding to each group (i.e., control or intervention) computed by the LMM model. Despite the slightly higher steps in the intervention group, the daily steps of the two groups did not differ substantially in the first 5 weeks. However, in the last 3 weeks, the intervention group had an average increase of 1,000 steps and the control group had an average decrease of 2,000 steps. We suspect that we fail to see differences in the early weeks due to the initial stimulation of participating in a fitness program. As time went by, the excitement from participation cooled down and the impact of the BAA algorithm started to dominate. We further defined adherent users to be those who used the CalFit app for 80% of the days during the study period. Under this criterion, 2 of the 7 users in the control group and 1 of the 6 users in the intervention group were identified as non-adherent. However, the difference in adherence percentage was not statistically significant ($p = 0.61$) between the two groups, primarily due to the small sample size.

Results of Qualitative Interview

During the second in-person session at 10-weeks, a trained research staff member interviewed the users on their experience. All users agreed that the CalFit app was easy to navigate, required minimal effort on the user side, and the number of push notifications was about right. One user in the intervention group told us, "I am excited to know my step goal every morning! I know I am doing well if my goal increases, and I know I need to keep up when my goal decreases." Another user in the control group, however, stated, "The goals are always the same. It's impossible for me to get that many steps so I stopped tracking."

DISCUSSION

The mSTAR study reveals the potential of using personalized step goals to facilitate physical activity. Interestingly, users' daily steps did not increase at a constant rate over the 10-week period. Rather, we observe that the daily steps of the two groups did not differ significantly in the first 5 weeks. But in the last 3 weeks, the intervention group was taking many more steps than the control group. We believe that in the first several weeks, physical activity was driven by users' initial enthusiasm with the start of their participation in the study.

However, when this enthusiasm wore out after 5 weeks, we observed significant difference in physical activity behavior between the two groups, suggesting the potential of using the CalFit app (and its underlying features such as the automated generation of personalized step goals using reinforcement learning) to deliver physical activity interventions.

DESIGN IMPLICATIONS

There are two major challenges associated with providing fully automated smartphone-based physical activity interventions. The first challenge is supporting users through key behavior change features and effective goal-setting in order to increase their level of physical activity. The second challenge is ensuring sustained maintenance of any increases in physical activity initiated by an intervention. Typical physical activity interventions address these challenges through frequent in-person coaching sessions, which are effective in initiating and maintaining behavior change. Since in-person coaching is expensive, mobile physical activity interventions seek to lower costs by reducing the amount of coaching. As a result, meeting these two challenges is substantially more difficult for fully automated smartphone-based physical activity interventions.

The mSTAR study demonstrates the potential of adopting behavior-change features and using personalization in mobile physical activity interventions to address these challenges. In particular, we found that sending one or two push notifications serves as a useful reminder. Furthermore, users prefer apps that do not require too much time and effort. Features that require regular user input, such as setting personal goals or keeping a diary to record steps/food intake, can create a burden on app adherence. Another main design choice is personalization. The BAA algorithm that sets personalized step goals for users is shown to be effective in increasing daily steps. Providing challenging but yet attainable goals can induce goal-achieving incentives, and giving daily feedback on performance (i.e., reminder push notification on daily goal and congratulating push notification) further reinforces exercise motivation. Conversely, fixed steps goals (10,000 steps/day) with no personalization can be unrealistically high or too easy to achieve and hinders users from progressing to be active.

Future designs of mobile fitness apps should consider personalized interventions, including but not limited to goals, push notifications, and displays. In addition, algorithms for goal-setting should take the complete history of the user as the basis to generate future interventions, particularly when the input and target metrics have high day-to-day variations. Implementing behavior change features, such as self-monitoring and summary feedback on performance, can further motivate physical activity. Overall, the app should be easy to navigate and require minimum manual inputs from users, particularly by using algorithms to automate personalization.

LIMITATIONS AND FUTURE WORK

One limitation of our study is the relatively small sample size. A larger scale study should be performed to further confirm the findings. In addition, the population of the study is university students, who may not be as concerned about their physical wellness as other populations (i.e., middle aged and elderly adults managing their chronic diseases). Another limitation is that the study lasted for only 10-weeks, so the long-term impact of the CalFit app is unclear.

In the future, we would like to extend our observations further by studying hypotheses in three directions. Firstly, how do different goal setting sources (i.e, self-set, trainer-set, and machine set) impact the intervention outcome? Secondly, how do different dynamic goal setting algorithms impact the intervention outcome? In particular, it would be beneficial to unveil if the success of this study is due to the BAA algorithm or due to the fact that step goals are not steady. We would like to compare the BAA algorithm to simpler analytical algorithms, such as, for example, setting the goal to be the 60th percentile of the steps in the past week. Thirdly, we would like to isolate the impact of the various design features (i.e., push notification, history tab, etc.) to provide recommendations on the most effective features to future fitness app designers.

CONCLUSION

We developed a novel fitness app called CalFit to track and deliver physical activity interventions. The app implements a reinforcement learning algorithm adapted to the context of generating personalized and adaptive daily step goals for each user so that the goals are challenging but attainable. Furthermore, the app adopts many behavior change features such as self-monitoring and customized feedback. A pilot study with 13 university students demonstrated that setting personalized step goals resulted in 2,200 more daily steps than setting steady step goals (of 10,000 steps/day) after 10 weeks. We believe the CalFit app (and its underlying features like the automated generation of personalized step goals using reinforcement learning) has the potential to deliver physical activity interventions in a fully automated fashion. A large scale, randomized controlled trial of a fully automated physical activity intervention is warranted.

ACKNOWLEDGMENTS

The authors would like to thank Emily Ma, Smita Jain, and Jessica Lin for their help during the pre-study in-person session. The authors would also like to thank Mingyang Li for his help in developing the app. This study was supported in part by funding from the Philippine-California Advanced Research Institutes (PCARI), funding from the UC Center for Information Technology Research in the Interest of Society (CITRIS), the Philippine-California Advanced Research Institutes (PCARI) grant IIID-2015-07, a grant (K24NR015812) from the National Institute of Nursing Research, and a grant from the National Center for Advancing Translational Sciences of the National Institutes of Health (KL2TR000143).

REFERENCES

1. Marc A Adams, James F Sallis, Gregory J Norman, Melbourne F Hovell, Eric B Hekler, and Elyse Perata. 2013. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. *PloS one* 8, 12 (2013), e82901.

2. Tim Althoff, Rok Sosič, Jennifer L Hicks, Abby C King, Scott L Delp, and Jure Leskovec. 2017. Large-scale

physical activity data reveal worldwide activity inequality. *Nature* 547, 7663 (2017), 336–339.

3. James J Annesi. 2002. Goal-setting protocol in adherence to exercise by Italian adults. *Perceptual and motor skills* 94, 2 (2002), 453–458.
4. MOSEK ApS. 2016. *The MOSEK optimization toolbox for MATLAB manual. Version 7.1.* `http://docs.mosek.com/7.1/toolbox/index.html`
5. Anil Aswani, Philip Kaminsky, Yonatan Mintz, Elena Flowers, and Yoshimi Fukuoka. 2016. Behavioral modeling in weight loss interventions. (2016). `https://ssrn.com/abstract=2838443`
6. Anil Aswani, Zuo-Jun Max Shen, and Auyon Siddiq. 2018. Inverse optimization with noisy data. *Operations Research* (2018), To appear.
7. Audie A Atienza, Abby C King, Brian M Oliveira, David K Ahn, and Christopher D Gardner. 2008. Using hand-held computer technologies to improve dietary intake. *American journal of preventive medicine* 34, 6 (2008), 514–518.
8. Kristen MJ Azar, Lenard I Lesser, Brian Y Laing, Janna Stephens, Magi S Aurora, Lora E Burke, and Latha P Palaniappan. 2013. Mobile applications for weight management: theory-based content analysis. *American journal of preventive medicine* 45, 5 (2013), 583–589.
9. Albert Bandura. 1991. Social cognitive theory of moral thought and action. *Handbook of moral behavior and development* 1 (1991), 45–103.
10. Marco Bardus, Samantha B van Beurden, Jane R Smith, and Charles Abraham. 2016. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *International Journal of Behavioral Nutrition and Physical Activity* 13, 1 (2016), 35.
11. Stephanie Bauer, Judith de Niet, Reinier Timman, and Hans Kordy. 2010. Enhancement of care through self-monitoring and tailored feedback via text messaging and their use in the treatment of childhood overweight. *Patient education and counseling* 79, 3 (2010), 315–319.
12. Jeannette M Beasley, William T Riley, Amanda Davis, and Jatinder Singh. 2008. Evaluation of a PDA-based dietary assessment and intervention program: a randomized controlled trial. *Journal of the American College of Nutrition* 27, 2 (2008), 280–286.
13. Lora E Burke, Molly B Conroy, Susan M Sereika, Okan U Elci, Mindi A Styn, Sushama D Acharya, Mary A Sevick, Linda J Ewing, and Karen Glanz. 2011. The effect of electronic self-monitoring on weight loss and dietary intake: a randomized behavioral weight loss trial. *Obesity* 19, 2 (2011), 338–344.
14. Meredith A Case, Holland A Burwick, Kevin G Volpp, and Mitesh S Patel. 2015. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA* 313, 6 (2015), 625–626.
15. Centers for Disease Control and Prevention. 2017. *Exercise or Physical Activity*. Technical Report.
16. Sunny Consolvo, Predrag Klasnja, David W McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A Landay. 2008. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 54–63.
17. Brianna S Fjeldsoe, Yvette D Miller, and Alison L Marshall. 2010. MobileMums: a randomized controlled trial of an SMS-based physical activity intervention. *Annals of Behavioral Medicine* 39, 2 (2010), 101–111.
18. Yuichi Fujiki. 2010. iPhone as a physical activity measurement platform. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. ACM, 4315–4320.
19. Yoshimi Fukuoka, Caryl L Gay, Kevin L Joiner, and Eric Vittinghoff. 2015. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. *American journal of preventive medicine* 49, 2 (2015), 223–237.
20. Yoshimi Fukuoka, Eric Vittinghoff, So Son Jong, and William Haskell. 2010. Innovation to motivation – pilot study of a mobile phone intervention to increase physical activity among sedentary women. *Preventive medicine* 51, 3 (2010), 287–289.
21. Dominik Gall, Jean-Luc Lugrin, Dennis Wiebusch, and Marc Erich Latoschik. 2016. Remind Me: An Adaptive Recommendation-Based Simulation of Biographic Associations. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*. ACM, 191–195.
22. Gurobi Optimization, Inc. 2016. Gurobi Optimizer Reference Manual. (2016). `http://www.gurobi.com`
23. Irja Haapala, Noël C Barengo, Simon Biggs, Leena Surakka, and Pirjo Manninen. 2009. Weight loss by mobile phone: a 1-year effectiveness study. *Public health nutrition* 12, 12 (2009), 2382–2391.
24. Dylan Hadfield-Menell, Stuart J Russell, Pieter Abbeel, and Anca Dragan. 2016. Cooperative inverse reinforcement learning. In *Advances in neural information processing systems*. 3909–3917.
25. Eric B Hekler, Matthew P Buman, Lauren Grieco, Mary Rosenberger, Sandra J Winter, William Haskell, and Abby C King. 2015. Validation of physical activity tracking via android smartphones compared to ActiGraph accelerometer: laboratory-based and free-living validation studies. *JMIR mHealth and uHealth* 3, 2 (2015).
26. Robert Hurling, Michael Catt, Marco De Boni, Bruce William Fairley, Tina Hurst, Peter Murray, Alannah Richardson, and Jaspreet Singh Sodhi. 2007. Using internet and mobile phone technology to deliver an

automated physical activity program: randomized controlled trial. *Journal of medical Internet research* 9, 2 (2007).

27. IBM. 2016. IBM ILOG CPLEX Optimization Studio. (2016).
28. John M Jakicic, Kelliann K Davis, Renee J Rogers, Wendy C King, Marsha D Marcus, Diane Helsel, Amy D Rickman, Abdus S Wahed, and Steven H Belle. 2016. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. *Jama* 316, 11 (2016), 1161–1171.
29. Nam-Seok Joo and Bom-Taeck Kim. 2007. Mobile phone short message service messaging for behaviour modification in a community-based weight control programme in Korea. *Journal of Telemedicine and Telecare* 13, 8 (2007), 416–420.
30. Rosemary Josekutty Thomas, Judith Masthoff, and Nir Oren. 2017. Personalising Healthy Eating Messages to Age, Gender and Personality: Using Cialdini's Principles and Framing. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*. ACM, 81–84.
31. Julie J Keysor. 2003. Does late-life physical activity or exercise prevent or minimize disablement?: a critical review of the scientific evidence. *American journal of preventive medicine* 25, 3 (2003), 129–136.
32. Abby C King, David K Ahn, Brian M Oliveira, Audie A Atienza, Cynthia M Castro, and Christopher D Gardner. 2008. Promoting physical activity through hand-held computer technology. *American journal of preventive medicine* 34, 2 (2008), 138–142.
33. Joseph A Knight. 2012. Physical inactivity: associated diseases and disorders. *Annals of Clinical & Laboratory Science* 42, 3 (2012), 320–337.
34. Artie Konrad, Victoria Bellotti, Nicole Crenshaw, Simon Tucker, Les Nelson, Honglu Du, Peter Pirolli, and Steve Whittaker. 2015. Finding the adaptive sweet spot: Balancing compliance and achievement in automated stress reduction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3829–3838.
35. Sanjay Krishnan, Animesh Garg, Richard Liaw, Lauren Miller, Florian T Pokorny, and Ken Goldberg. 2016. Hirl: Hierarchical inverse reinforcement learning for long-horizon tasks with delayed rewards. (2016). `http://arxiv.org/abs/1604.06508`
36. Hyunho Lee and Youngseok Lee. 2017. A Look at Wearable Abandonment. In *Mobile Data Management (MDM), 2017 18th IEEE International Conference on*. IEEE, 392–393.
37. Edwin A Locke and Gary P Latham. 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist* 57, 9 (2002), 705.
38. Edward McAuley and Bryan Blissmer. 2000. Self-efficacy determinants and consequences of physical activity. *Exerc Sport Sci Rev* 28, 2 (2000), 85–88.
39. Kathryn Mercer, Melissa Li, Lora Giangregorio, Catherine Burns, and Kelly Grindrod. 2016. Behavior change techniques present in wearable activity trackers: a critical analysis. *JMIR mHealth and uHealth* 4, 2 (2016).
40. Yonatan Mintz, Anil Aswani, Philip Kaminsky, Elena Flowers, and Yoshimi Fukuoka. 2017. Behavioral Analytics for Myopic Agents. (2017). `http://arxiv.org/abs/1702.05496`
41. Sean A Munson and Sunny Consolvo. 2012. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In *Pervasive computing technologies for healthcare (PervasiveHealth), 2012 6th international conference on*. IEEE, 25–32.
42. Andrew Y Ng, Stuart J Russell, and others. 2000. Algorithms for inverse reinforcement learning.. In *Icml*. 663–670.
43. Parse. 2016. Parse Documentation. (2016). `http://docs.parseplatform.org`
44. Kevin Patrick, Fred Raab, Marc A Adams, Lindsay Dillon, Marian Zabinski, Cheryl L Rock, William G Griswold, and Gregory J Norman. 2009. A text message–based intervention for weight loss: randomized controlled trial. *Journal of medical Internet research* 11, 1 (2009).
45. Pew Research Center. 2016. *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies*. Technical Report.
46. William T Riley, Daniel E Rivera, Audie A Atienza, Wendy Nilsen, Susannah M Allison, and Robin Mermelstein. 2011. Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational behavioral medicine* 1, 1 (2011), 53–71.
47. Sarah Knapton. 2017. The 10,000 Steps a Day Myth: How Fitness Apps Can Do More Harm Than Good. (2017). `http://www.telegraph.co.uk/news/2017/02/21/10000-steps-day-myth-fitness-apps-can-do-harm-good/`
48. Oliver S Schneider, Karon E MacLean, Kerem Altun, Idin Karuei, and Michael Wu. 2013. Real-time gait classification for persuasive smartphone apps: structuring the literature and pushing the limits. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 161–172.
49. Richard S Sutton and Andrew G Barto. 1998. *Reinforcement learning: An introduction*. MIT press Cambridge.
50. Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. 2000. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*. 1057–1063.

51. U.S. Department of Health and Human Services. 2008. Physical Activity Guidelines for Americans. (2008). https://health.gov/paguidelines/pdf/paguide.pdf
52. Youri van Pinxteren, Gijs Geleijnse, and Paul Kamsteeg. 2011. Deriving a recipe similarity measure for recommending healthful meals. In *Proceedings of the 16th international conference on Intelligent user interfaces*. ACM, 105–114.
53. Ted Vickey, John Breslin, and Antonio Williams. 2012. Fitness–There's an App for That: Review of Mobile Fitness Apps. *International Journal of Sport & Society* 3, 4 (2012).
54. World Health Organization. 2017a. Health topics: physical activity. (2017). http://www.who.int/topics/physical_activity/en/
55. World Health Organization. 2017b. Physical activity fact sheet. (2017). http://www.who.int/mediacentre/factsheets/fs385/en/
56. World Heart Federation. 2017. Physical inactivity. (2017). http://www.world-heart-federation.org/cardiovascular-health/cardiovascular-disease-risk-factors/physical-inactivity/