HHS Public Access
Author manuscript
J Behav Med. Author manuscript; available in PMC 2019 May 17.

Published in final edited form as:
J Behav Med. 2019 April ; 42(2): 276–290. doi:10.1007/s10865-018-9964-1.

Can the artificial intelligence technique of reinforcement learning use continuously-monitored digital data to optimize treatment for weight loss?

Evan M. Forman[1,2], **Stephanie G. Kerrigan**[1,2], **Meghan L. Butryn**[1,2], **Adrienne S. Juarascio**[1,2], **Stephanie M. Manasse**[2], **Santiago Ontañón**[3], **Diane H. Dallal**[1,2], **Rebecca J. Crochiere**[1,2], and **Danielle Moskow**[1,2]

[1]Department of Psychology, WELL Center, Drexel University, Stratton Hall, 3141 Chestnut Street, Philadelphia, PA 19104, USA

[2]Center for Weight, Eating, and Lifestyle Sciences (WELL Center), Drexel University, 3201 Chestnut Street, Philadelphia, PA 19104, USA

[3]Department of Computer Science, Drexel University, 3401 Market Street, Philadelphia, PA 19104, USA

Abstract

Behavioral weight loss (WL) trials show that, on average, participants regain lost weight unless provided long-term, intensive—and thus costly—intervention. Optimization solutions have shown mixed success. The artificial intelligence principle of “reinforcement learning” (RL) offers a new and more sophisticated form of optimization in which the intensity of each individual’s intervention is continuously adjusted depending on patterns of response. In this pilot, we evaluated the feasibility and acceptability of a RL-based WL intervention, and whether optimization would achieve equivalent benefit at a reduced cost compared to a non-optimized intensive intervention. Participants ($n = 52$) completed a 1-month, group-based in-person behavioral WL intervention and then (in Phase II) were randomly assigned to receive 3 months of twice-weekly remote interventions that were non-optimized (NO; 10-min phone calls) or optimized (a combination of phone calls, text exchanges, and automated messages selected by an algorithm). The Individually-Optimized (IO) and Group-Optimized (GO) algorithms selected interventions based on past performance of each intervention for each participant, and for each group member that fit into a fixed amount of time (e.g., 1 h), respectively. Results indicated that the system was feasible to deploy and acceptable to participants and coaches. As hypothesized, we were able to achieve equivalent Phase II weight losses (NO = 4.42%, IO = 4.56%, GO = 4.39%) at roughly one-third the cost (1.73 and 1.77 coaching hours/participant for IO and GO, versus 4.38 for NO), indicating strong promise for a RL system approach to weight loss and maintenance.

Compliance with ethical standards

Conflict of interest Evan M. Forman, Stephanie G. Kerrigan, Meghan L. Butryn, Adrienne S. Juarascio, Stephanie M. Manasse, Santiago Ontañón, Diane H. Dallal, Rebecca J. Crochiere and Danielle Moskow declare that they have no conflict of interest.

Human and animal rights and Informed consent All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors. Informed consent was obtained from all individual participants included in the study.

Keywords

Artificial intelligence; Behavioral treatment; Lifestyle modification; Weight loss; Obesity; Optimization

Introduction

More than 70% of Americans are overweight (Centers for Disease Control and Prevention (CDC), 2014); among these, 49% are currently attempting to follow a weight loss program (Snook et al., 2017), with even more having previously made such attempts. Individuals attempting weight loss do so through many methods including clinician-guided weight loss interventions, commercial weight loss programs and self-help programs, and mobile health weight loss applications (apps). Collectively, empirical data from these interventions suggest that weight lost is typically regained unless some form of supportive accountability (e.g., in-person or telephone coaching) is provided and is sustained long-term (Anderson et al., 1999; Kramer et al., 1989; Wadden et al., 1989; Wilson, 1994; Wilson & Brownell, 2002). However, providing such coaching is expensive (Diabetes Prevention Program Research Group, 2003; Jeffery et al., 2000; Krukowski et al., 2011), requires expert clinicians, of which there is a shortage (Association of Behavioral and Cognitive Therapies, 2018; Society of Clinical Psychology, 2018), and depends on participants tolerating the intensive time and scheduling commitment. Automated coaching is an attractive alternative because it is low-cost and scalable. However, weight losses produced by automated coaching are quite small, on average (i.e., 1–3 kg), probably because they are unable to provide accountability in the same way that human coaching can (Joo & Kim, 2007; McGraa, 2010; Patrick et al., 2009).

Of note, variability of weight loss is high, both in low-intensity and high-intensity programs. For example, while the mean weight loss of low-intensity interventions is low, a substantial subset (e.g., ~ 23–47%) of those in low-intensity interventions reach clinically significant weight loss (i.e., ≥ 5%) (Brindal et al., 2012; Webber et al., 2008). Conversely, while higher-intensity interventions produce average weight losses of approximately 10% (Wadden et al., 2007), about one-fourth of participants fail to reach a clinically significant weight loss even after a year or more (Butryn et al., 2017, 2018; Forman et al., 2016).

An innovative solution to the trade-off between cost and accountability would be a system that can adapt to the needs of the individual by allocating resources away from those unable to benefit from them and towards those most able to benefit. One system for doing so, i.e., stepped care, measures weight loss at an established point, and “steps up” or “steps down” the intensity (i.e., frequency) of intervention depending on whether weight losses are below or above expected targets. By adapting to an individual’s observed needs and conserving resources where they are not needed, outcomes can be enhanced and costs can be reduced. Stepped care interventions for weight loss have shown initial promise in producing superior weight loss, increased physical activity, dietary improvements, and cost savings compared to standard behavioral treatments (Carels et al., 2005, 2007; Jakicic et al., 2012). Importantly, they have also demonstrated that a subset of participants achieve weight loss benchmarks even when provided with minimal support (Carels et al., 2017). However, several studies have observed significantly less weight loss in participants who were stepped down to less intensive interventions after initial success (Carels et al., 2013; Jakicic et al., 2012), suggesting that the reduction was too steep or should not have happened at all for many initial responders. Moreover, in a two-step study, a majority of participants who were stepped-up twice continued to fail (Carels et al., 2017), calling into question whether continuation of an intensive (but ineffective) treatment program over an 18 month period is an effective use of limited resources and whether the provided interventions are of the optimal modality for this subset of patients. These findings point to several related shortcomings of stepped care systems. First, few intensity adjustments can be made; yet, people's needs vary through time. Second, each step up or down is, by necessity, relatively large, which compromises effectiveness. Third, assignment is based on whether participants' weight loss meets an overall, pre-established group standard and not a relative, individualized standard. Thus, participants who do reasonably well by group standards but who could benefit exponentially from a higher level of intensity will never have that opportunity. Moreover, participants who do poorly at a lower level of intensity will be stepped up to a high level of care, and will continue to receive the highest-cost services even if they do just as poorly in this higher level of care. As such, potential cost savings are not realized.

"Control systems" interventions represent another type of adaptive intervention. These interventions use theory about change mechanisms (e.g., "ambitious but doable" goals will maximize behavior change) to specify an algorithm that adapts the intervention depending on relevant continuously-collected data. An example of such a system was recently deployed that dynamically alters physical activity goals. The adaptive intervention was administered on a smartphone application, in which new physical activity goals were continuously assigned and positive reinforcement was provided for meeting goals. While using the application, participants increased their daily steps by over 2600 per day, as compared to the baseline period in which physical activity goals were not specified (Korinek et al., 2018). Additionally, individual differences in predictors of success on a given day support that an individual-level of tailoring in interventions is justified and may further improve outcomes (Phatak et al., 2018). Of note, control systems interventions do not have the capacity to adjust for specific participants who might not be responding well to a particular type of intervention (e.g. "ambitious").

The artificial intelligence (AI) strategy of "reinforcement learning" (Sutton & Barto, 1998) is an especially novel adaptive solution. Whereas step-up and step-down interventions normally choose a single switchover point, reinforcement learning continuously monitors outcomes and adjusts actions (e.g., type of intervention delivered) throughout the intervention period. As such, reinforcement learning is able to optimize intervention selection to a more precise degree than can step-up/step-down interventions. Moreover, unlike control systems, the optimization component of reinforcement learning is not reliant on a (potentially flawed) theoretical model but instead is driven entirely by specified outcomes. Reinforcement learning optimizes intervention choice by comparing interventions within individual, i.e., choosing the best intervention for each individual, relative to all other interventions for that individual. Essentially, an intervention is repeated many times, and on each occasion, a *reward score* is calculated for each intervention, i.e., a quantification or computational sum of all relevant outcomes attributable to the previous intervention. An average reward score is maintained for each intervention for each person, and serves as an index of the relative efficacy of each intervention for each person. The intervention with the highest average reward score for a given person is then delivered to the individual, which is termed an *exploit* strategy. If the efficacy of the intervention decreases over time, and the average reward score of a different intervention rises to the top, this second intervention will be delivered instead. (The term *reinforcement learning* comes from the parallel with operant conditioning, i.e., the AI system "learns" the best actions by repeatedly trying out actions and then calculating and tracking the quantified consequences of those actions.) In order to allow for changing conditions that might alter the relative efficacy of interventions during the course of treatment, the remaining interventions are occasionally *explored* (i.e., delivered instead of the intervention that would have been selected by the *exploit* strategy). AI researchers have developed efficient optimization algorithms for determining the best balance between *exploring* and *exploiting* (Auer et al., 2002). Importantly, because a reinforcement learning system uses intra-individual comparisons and can incorporate intervention costs, it can be used to derive the best performance at the lowest cost. In fact, a *cost-containment strategy* can be used to optimize outcomes within the confines of a pre-set, fixed cost. Figure 1 represents a visual depiction of a reinforcement learning system's use of the exploit and explore strategies to repeatedly select interventions for each participant.

While the above description is written as if applied to behavioral interventions, reinforcement learning was initially developed for, and has shown remarkable successes in, non-person-based engineering cases. Examples include optimization of manufacturing processes (Aydin & Öztemel, 2000), automation of driving and helicopter piloting systems (Abbeel & Ng, 2004; Kuderer et al., 2015; Ng et al., 2006) and development of artificial intelligence-driven strategy board game systems including one that can now beat the world's foremost players in the ancient Chinese game of Go (Silver et al., 2016).

In fact, only a few examples of reinforcement learning have been reported in the health behavior fields. One example is *MyBehavior*, a smartphone app that delivered personalized interventions for promoting physical activity and dietary health, tracked outcomes after each intervention, combined them to form a reward score, and then used average reward scores to select the next intervention (Rabbi et al., 2015). While using *MyBehavior*, participants demonstrated greater 14-week improvement in PA and calorie intake, compared to the initial control phase (in which participants received automated, generic prescriptive recommendations for diet and physical activity from a pool pre-generated by health experts). Participants also reported *MyBehavior*'s suggestions to be more relevant to their lives and endorsed a greater desire to follow the personalized suggestions. Another study reported that a reinforcement learning system that varied step goals produced greater physical activity and generated more challenging but attainable step goals than a non-tailored step goal control (Zhou et al., 2018). A similar study used reinforcement learning to vary text message content and produced better physical activity and blood glucose levels in diabetic adults, compared to non-optimized, weekly physical activity reminders (Yom-Tov et al., 2017). Finally, reinforcement learning proved effective in reducing momentary stress, compared to a non-optimized control with randomly recommended interventions, by selecting mobile phone-based stress management techniques based on patterns of previous response (Paredes et al., 2014).

Of note, two simulation studies used previously-collected data to support the use of reinforcement learning. One study suggested that using reinforcement learning to adapt the content and frequency of text messages would have substantially improved medication adherence in chronically ill patients (Piette et al., 2014). Compared to common approaches of improving adherence through text messages (e.g., patient reminders, tailoring based on a baseline survey), the reinforcement learning system was estimated to better individualize messages based on the underlying causes of non-adherence, adapt to changes in those underlying causes, and tailor messaging frequency to prevent feedback fatigue. A second simulation study indicated that using reinforcement learning to repeatedly vary the intensity of an intervention (interactive voice call versus short coaching call versus long coaching call) for patients with chronic pain would achieve equivalent improvements in physical activity to an hour-long phone call for all patients, but would only use 44% of the clinician time (Piette et al., 2016).

Reinforcement learning could be an especially efficient method of optimizing the amount of human coaching provided in a weight loss intervention such that the minimum resources are expended to produce outcomes equivalent to a high level of coaching intervention (thus saving intervention costs). In particular, the current pilot study evaluated a reinforcement learning system that could select between three intensities of coaching: (1) fully automated coaching that is delivered through computer-generated texts that are tailored to participant outcome data, (2) low-intensity coaching in the form of short text exchanges and (3) high-intensity coaching in the form of phone calls. In this study, several outcomes were continuously monitored: frequency of self-monitoring of eating and weight, frequency of meeting calorie goals and physical activity goals, and amount of weight loss. These outcome data were used to create a reward score for each of the three interventions for each participant, and the likelihood of an intervention occurring for a given individual was determined by its reward score. One practical downside of having patient interventions determined solely by reward score is that coaches have no way of anticipating the total amount of coaching time that will be required on a given day. Thus, the current study also tested a modified reinforcement learning system in which intervention assignments were constrained by a fixed amount of daily coaching time.

As such, the current study enrolled 52 participants in a 16-week weight loss program across two phases. In the first phase, all participants received a 4-week, weekly, group-based, in-person behavioral weight loss program. In the second phase, participants were randomly assigned to one of three types of 12-week, twice-weekly remote coaching interventions: non-optimized (i.e., all participants always receive high-intensity coaching), or one of two AI-optimized interventions, including individually-optimized (i.e., at each of the 24 intervention points, participants receive the intervention with the highest reward score for them so far, except when the system is “exploring”) or group-optimized (i.e., interventions are assigned based on the highest possible *total* reward scores, across all interventions assigned, given a predetermined amount of total intervention time across all participants for the day). The aims of this pilot study were to evaluate the intervention’s feasibility and acceptability, characterize intervention selection between individuals and across time, within individual, and establish whether either form of AI optimization produces equivalent weight losses at lower cost.

Methods

Participants

Participants (*N* = 52) were deemed eligible for inclusion if they had a body mass index (BMI) between 25 and 50 kg/m^2, were between 18 and 70 years of age and had a smartphone capable of receiving text messages and with Bluetooth connectivity. Exclusion criteria were inability to engage in physical activity, using weight-affecting medication, pregnant or planned to become pregnant during the study period, had lost more than 5% of their weight in the past 6 months, or met criteria for binge eating disorder.

Procedure

The study protocol was approved by the Drexel University Institutional Review Board. Participants were recruited through advertisements on local radio stations and on Facebook. Preliminary eligibility was established through an online survey and/or phone screen, with eligible and interested participants continuing to an in-person baseline assessment. During enrollment, all aspects of study procedures were explained including the need to purchase and use daily a Fitbit Flex 2 activity tracker, Yunmai Smartscale wireless scale and Fitbit app, and to provide a sharing token such that data recorded would be synced to study servers. Participants were asked if the combined cost (approximately $110) of the activity tracker and scale represented a financial hardship and, if so, these items were loaned by the study. For the 16 weeks of in-person and remote coaching, participants were instructed to wear the activity tracker during all waking hours, weigh themselves each morning on the wireless scale, and record everything they ate and drank in the Fitbit app. Participants were compensated $150 for completing assessments.

Once enrolled, participants attended 4 weeks of in-person behavioral weight loss group sessions, led by a coach with behavioral weight loss expertise. At the completion of the in-person phase (Phase I), participants were randomly assigned to one of three versions of a 12-week remote coaching (Phase II): non optimized, individually optimized or group optimized (see below). Randomization was blocked by weight lost at 4 weeks and used a 1:2:2 ratio favoring the two optimization conditions (in order to better assess characteristics of the novel intervention system).

Intervention phases

Phase I: In-person groups—The in-person phase of intervention consisted of four weekly, closed group sessions of behavioral weight loss. This phase was considered necessary for participants to adequately learn foundational nutrition and behavioral weight loss strategies and to establish a relationship with their coach. Groups lasted 90 min, consisted of 16–18 participants, were manualized, and consisted of brief check-ins on each individual's progress, skill presentation, and a skill building exercise. Content was based on Look AHEAD and the Diabetes Prevention Program protocols and included nutritional education, self-monitoring of calorie intake, goal setting, stimulus control, social support, problem solving and relapse prevention. Prescriptions were set for a balanced-deficit diet with a goal of limiting intake to 1200–1800 kcal/day (depending on initial weight). The physical activity prescription gradually increased to 30 min of moderate-vigorous activity per day by week 4.

Phase II: Remote coaching—Remote coaching took the form of twice weekly text messages and/or phone calls. As per Table 1, three types of remote coaching were delivered depending on study condition and optimization parameters: (1) Automated text message (i.e., a computer-generated message tailored based on behavioral, and weight data), (2) Coach text exchange (i.e., up to 3 coach and 3 participant messages about behavioral adherence, problem-solving, and/or planning behavioral strategies for the upcoming days before the next coaching session), and (3) Coach phone call (i.e., a 12-min phone call focusing on checking in about progress towards goals, reviewing strategies discussed during the 4-week workshop, praising behavioral adherence, problem-solving inadherence, and/or planning behavioral strategies for the upcoming days before the next coaching session.)

Participants were told to keep a predetermined hour slot of time open twice a week, at the same time on those 2 days, for a total of 24 coaching sessions in the remote coaching phase. The coach utilized a portal that logged all participant information and integrated an algorithm that would convey who received what intervention each day.

Optimization conditions

As per Table 2, participants were randomized to one of three conditions. In the *Non-Optimized Condition* all participants received twice-weekly coach phone calls. In the *Individually Optimized Condition* participants were assigned, every 3–4 days, the intervention with the highest "reward score" for that participant, adjusted by intervention cost (see below). In the *Group Optimized Condition* participants were assigned the interventions that maximized "reward score" for the group as a whole and could fit within a certain amount of time (time available was equivalent to the number of participants multiplied by 5 min). For example, within one group of twelve participants, 4 auto-texts (0 min), 4 text exchanges (12 min) and 4 phone calls (48 min) might be assigned. Of note, a given individual might be assigned an intervention with that has a lower score than an alternate intervention if doing so maximizes that reward score total for the group (by allowing that time to be used for another group members with a higher discrepancy between reward scores).

Reinforcement learning optimization

Calculation of reward score—In the remote intervention phase, at the end of every 3–4-day intervention period, a reward score was calculated for each participant and assigned to the intervention most recently provided. The individual reward score was calculated as: the sum of days (1) self-monitored weight (i.e., wireless scale provided a valid weight), (2) days self-monitored eating (i.e., defined as entering ≥ 3 foods and ≥ 300 calories), (3) days met calorie goal (which could be adjusted in collaboration with coach), and (4) days met PA goal, and (5) weight loss in pounds (such that weight losses increase the reward score and weight gains decrease it), divided by the number of days in that intervention period (to account for the differences between 3- and 4-day intervention periods). The *average* reward score (i.e., running average of all rewards scores for that intervention for that participant) for that intervention was then updated.

Reward score adjustment—In the individually optimized condition, the reward score was adjusted for intervention cost (i.e., time) in each group so that the chosen reward balanced cost with need (e.g., if a very small amount differentiated the reward score for a phone call and an automated text, the need for the phone call is not worth the cost of the phone call). The reward score for each intervention was divided by a cost assigned to that intervention. At the start of the remote intervention, costs were set to: automated text: 1; text exchange: 1.1; phone call: 1.3; they were subsequently adjusted to 1, 1.1 and 1.11, respectively based on observed time cost of text exchanges compared to phone calls and intervention selection frequency.

Exploit versus explore schedules—Optimization was primarily based on an “exploit” strategy. Twice each week, the AI system “exploited” its knowledge by selecting the intervention with the highest cumulative cost-adjusted reward score for each individually-optimized participant. Exceptions were 10 pre-set “explorations” of alternate interventions (described below in further detail). In group optimization, the formula sought to optimize the cumulative reinforcement learning across the group, “exploiting” the interventions for each person based on the relative need of that person within the group. Specifically, group optimization was achieved using the UCB1 algorithm (Auer et al., 2002), which essentially identifies the combination of interventions that maximizes the reward scores for all participants, within a group. However, in order to allow for a full sampling of interventions and to allow for change over time, an “explore” strategy was also utilized in each condition. In the individually optimized condition, the explore strategy was assigned for the first 6 remote interventions and thereafter at the 13th and 14th intervention, and the 19th and 20th intervention. (Though traditional reinforcement learning calls for the explore strategy to be selected probabilistically, we performed numerical simulations which indicated that predefined exploit/explore cycles were able to more quickly adapt to change over time, given that we only had 24 intervention assignments total.) Table 3 demonstrates how one individually-optimized participant’s data were translated into intervention selections during both exploit and explore cycles. In the group optimized condition, based on previous work (Ontanón, 2013, 2017), the UCB1 algorithm was utilized to balance the need for “exploiting” the best intervention (thus far) with the need for “exploring” interventions that had not been explored for a long time for a particular person (Auer et al., 2002). Specifically, the likelihood that a lower-reward-score intervention would be “explored” for a participant was proportional to how many days it had been since the last time this intervention had been delivered to that particular person. The explore/exploit assignment was manually overridden and set to auto-text on several occasions when the intervention fell on a holiday, because the coach would not have been available. (Participants in the non-optimized condition received no intervention on these days.)

Measures

Weight—Participants self-measured weight at approximately the same time each morning using a standardized Yunmai wireless mini smart scale that connected to Bluetooth, accurate to 0.1 kg. Participants were instructed to weigh themselves without any clothes on. Each scale was linked to the user's Fitbit app and data appeared in the online user portal, which was available to the research coordinator and coach.

Physical activity—Physical activity was measured in minutes of moderate-to-vigorous physical activity (MVPA) using the Fitbit Flex or similar type of Fitbit, a consumer-grade wrist-worn activity tracker.

Calorie intake—Participants logged all food and beverages using the Fitbit mobile phone application. They used this tool to calculate calorie intake and facilitate meeting the calorie goal set during Phase I (and sometimes adjusted in Phase II).

Acceptability—Participants were asked in their post-treatment survey how important the following behaviors had been in meeting their weight goals: weighing yourself daily, tracking everything you eat each day, setting and meeting daily calorie goals, tracking your exercise, and setting and meeting exercise goals. Participants also rated how effective they thought the program was in helping them lose weight, how satisfied they were with the approach we used to help them lose weight, if they would recommend this program to another adult wishing to lose weight, how confident they are they will be able to keep the weight off after the program ends, and how they liked the frequency and types of contact they received.

Data analysis

Data were analyzed using SPSS version 23. Weight change between Phase I and Phase II was calculated as the difference between the weight on the day of the first in-person group and the weight on the first day of remote coaching. Weight at the end of the intervention was the last recorded weight for each person. One individual discontinued the intervention (and self-weighing) prior to intervention end; thus, this technique used a last observation carried forward approach.

Results

Participant characteristics

Participants' mean age was 56.60 ± 13.35 years old. Mean BMI was 34.30 ± 5.76 kg/m^2. The sample was 67.3% female and 80.5% White, 15.7% African-American, 2.0% Asian, and 2.0% more than one race. On average, during Phase I of treatment, participants lost 2.78 ± 1.90% of their starting body weight. As per Table 4, participants showed general equivalence across treatment conditions.

Feasibility

The feasibility of the reinforcement learning intervention system was evaluated in three domains: the ability of the system to assign interventions based on the reward score, the ability of the coaches to use the system to deploy the interventions, and the ability of participants to receive and make use of the interventions. The technology underlying the intervention involved retrieving data from several inputs in participants' possession (i.e., wireless scale, activity tracker and food tracking app), which was successfully pulled onto a custom-created cloud-based application. For the most part, this application also successfully calculated reward scores, assigned interventions and tracked intervention assignment throughout the study. However, we discovered too late that the reward score average did not utilize a decay function as planned. As such, intervention assignment did not reflect our intention to make recent outcomes more impactful than less recent outcomes. The application did successfully generate text messages that were automatically tailored based on the participants' data. Part of the application consisted of an online user portal that allowed coaches to see which interventions had been assigned, access participants' data (including specific foods tracked, and graphs of calorie totals, weight, and physical activity) and text message with participants. Coaches were able to access information via the portal and to attempt all assigned interventions throughout the study. According to a short survey of coaches, the portal was easy to use ($M = 3.33$ out of 4) and that they were effectively able to carry out the remote coaching ($M = 3.33$ out of 4). However, Coaches noted several problems: text message exchanges took longer than expected, multiple ongoing text message threads were difficult to manage, the total intervention time for texts plus calls in the individually optimized group fluctuated more than was ideal, participants did not always respond to phone calls or text messages in a timely fashion (mid-way through the study, a text message reminder of the evening's intervention with a specific time to be available was implemented to improve participant responsiveness), some participants could sometimes benefit from a longer call to tackle complex barriers to weight control, and that there was less rapport built with participants for whom phone calls were rarely assigned.

Acceptability

Participants were asked specific questions regarding the interventions they received. Acceptability ratings were similar across conditions for all measures; thus, the results are presented collapsed across conditions. The majority (76.5%) reported that the contact frequency was satisfactory (2% reported wanting more frequent contact, 21.6% reported that twice per week was too much and wanted less). When provided a free-form prompt to discuss the most helpful pieces of the interventions, 34.6% specifically identified "accountability," "keeping me on track," or "support." When asked about the least helpful aspects of the phone calls, 21.2% discussed the timing of the calls or other commitments interfering with the phone call, 9.6% discussed the content of phone calls being unhelpful due to repetition of similar content call-to-call, and 42.3% reported that there was nothing they could list as unhelpful. Fewer individuals, 19.2%, identified that accountability was a helpful component of the text-message based interventions; 28.8% reported that the most helpful part of the text messages were that they served as a reminder and summarized recent progress. The most commonly reported "least helpful" part of the text messages were that the automated texts were impersonal, at times didn't reflect a personal situation (e.g., not self-weighing due to a vacation), or provided reinforcement when individuals felt they did not do well.

Across conditions, participants rated the program as effective for helping individuals lose weight ($M = 4.20 \pm 1.06$ out of 5) and were satisfied with the approach used (4.18 ± 0.82). Further, the low rate of attrition ($n = 1$ dropped out of the study during Phase II) evidences good acceptability of the intervention approach. Participants also reported being confident that they could maintain weight lost or continue to lose weight (3.71 ± 0.50) and nearly all participants stated they would recommend the program to another adult wishing to lose weight (72.5% would strongly recommend, 25.5% would recommend with some hesitation, 2% were not sure if they would recommend). Ratings of effectiveness and satisfaction were similar across conditions (with participants in the optimized conditions providing somewhat higher ratings).

Characterizing intervention selection

Across participants in the individually optimized and group optimized conditions, participants received approximately equal numbers of each intervention (Table 5). However, ranges were wide, indicating that each individual participant received a variable proportion of each intervention. The intervention with the highest reward score changed over the course of the intervention on average 2.3 times (maximum 6) for each individual (95% CI [1.76–2.84]).

Weight losses

Observed Phase II weight losses in non-optimized, individually optimized and group optimized were 4.42, 4.56 and 4.39%, respectively. Thus individually optimized exhibited slightly more weight loss than non-optimized, and the difference between non-optimized and group optimized (0.03) represented a near-zero effect size. The in-person and remote coaching combined produced 6.36, 7.50 and 7.08% weight loss across the three groups. Of note, about the same proportion of participants (58–70%) reached a 5% weight loss (a commonly-used cut-point indicating clinical significance) within each condition, with numbers highest in the optimized conditions.

Coach time spent on intervention delivery

On average during the remote phase of the intervention, coaches had 4.38 h of contact with each participant (all phone calls) in the non-optimized condition. Coaches had 1.73 and 1.77 h of contact with each participant in the individually optimized and group optimized conditions respectively through a mix of phone calls and text conversations. Figure 2 shows weight losses during the remote phase against time spent on intervening. Thus, the nonoptimized condition required approximately 2.5 times as much clinician time as the individually optimized and group optimized conditions for equivalent weight losses.

Adherence to behaviors

Each day, participants met an average of 2.79 ± 0.62 behavioral goals out of four (i.e., weighing self daily, recording food intake, meeting calorie goal, and meeting physical activity goal). Within each optimized condition, number of behavioral goals met based on the intervention most recently provided were similar between and within conditions (Table 6).

Discussion

This study characterized and evaluated the feasibility, acceptability, cost savings and effectiveness of an innovative AI optimization system for weight loss intervention called reinforcement learning which continuously optimized intervention selection for every participant based on an intervention scoring algorithm. Reinforcement learning has rarely been applied to the behavioral sciences and this is the first attempt to use its principles to optimize a behavioral weight loss program. While the system was mathematically and technically quite complicated, and some issues arose during the intervention that required problem-solving, we were able to successfully deploy it in a sample of 52 overweight participants who received 12 weeks of twice-weekly remote coaching following four in-person weight loss group sessions. The system was able to upload digital daily data on adherence to food logging, meeting calorie goals, meeting physical activity goals and weight change; combine these data mathematically into a reward score assigned to the most recent intervention; use the history of reward scores to assign one of three interventions to participants in an optimized condition; and send a tailored, automated text message (if indicated). The system was maintained within a web portal that coaches used to see assigned interventions and summarized data. Coaches reported that the system was relatively easy to use and that it worked well. A problem reported by coaches was that the total time taken up by the remote sessions varied a great deal from intervention day to intervention day across a group of participants assigned to Individual Optimization. As such, Group Optimization, which maximized group reward score within a (constant) time limit, may represent a more satisfactory and feasible optimization system in many contexts.

Participants also reported that they were satisfied with the system, no matter which condition they were assigned. Thus, participants reported no adverse reactions to having an overall reduced amount of contact with their weight loss coach or to having the intervention modality algorithmically determined, i.e., switch from time to time between phone call, text exchange, and auto-text. These types of changes are an unusual intervention feature within a lifestyle modification program, so participant satisfaction is significant. Previous work has observed that daily novelty in goals is acceptable and efficacious to participants as well (Korinek et al., 2018; Poirier et al., 2016); it may be that novelty in intervention components maintains participant engagement. Participants did express a desire for the automated text messages to be better personalized, so a future iteration of the system might employ more sophisticated tailoring. For example, the messaging could reflect changes coaches and participants' collaboratively made to goals, and could utilize more varied messages that adapt to participant response or increase the novelty of the automated messages.

Two of the principles behind reinforcement learning optimization are that (1) not everyone has the same intervention needs and (2) the intervention needs sometimes differ through time. To the extent that these assertions are true, cost savings can be extracted by delivering lower-cost interventions when doing so does not compromise outcomes. Our descriptive data suggest that, in fact, quite a lot of variability existed as to which interventions delivered the best response. For example, among optimized participants, phone calls were the optimal intervention (according to our algorithm) about 25% of the time, on average, but the individual range was from 8 to 50%. Similarly, automated text messages were the optimal intervention 17% of the time for some people but 75% for others, with an average of 45%. Moreover, it also turned out to be true that the intervention producing the optimal response varied *within* person, over time. In fact, the average number of times the optimal intervention changed within-person (according to our algorithm) was about 2 for individually optimized participants and about 3 for group optimized participants, with some participants experiencing as many as 6 changes. These changes, and the reinforcement learning system's ability to respond to them, represent potential improvements over step-down and step-up interventions. I.e., the reinforcement learning system was able to save costs during times when participants were in need of (or unable to make use of) less intensive interventions, and was possibly able to improve weight losses by shifting to a more intensive intervention when outcomes suggested the participant was no longer responding optimally to a less-intensive intervention. Step-up/down interventions often only make a single adjustment.

Perhaps most importantly, the reinforcement learning system was able to considerably lower costs in the optimized conditions while still producing equivalent weight losses. In fact, the non-optimized intervention cost 2.5 times as much as the optimized interventions, as measured by clinician time needed to deliver the remote interventions. Future work could attempt to derive even larger cost savings through improved optimization. For instance, further investigations could compare systems that use different formulas for computing reward score (e.g., emphasizing weight more than proximal behaviors, emphasizing behaviors historically difficult for each person, emphasizing behaviors that have historically best predicted longer-term weight loss), that apply intervention "costs" differently, that select which intervention is "optimal" differently, and that use different algorithms for determining when to explore and when to exploit. In addition, longer trials of reinforcement learning-based intervention systems would have the advantage of being able to capitalize on additional periods of data collection (allowing for more accurate reward scores) and testing the possibility that the effect of optimization could be greater the further in time the participant is from the more intensive in-person intervention.

Additionally, future work could measure baseline characteristics that are tied to the complexity of participants' coaching needs (e.g., depression and binge eating), based on the notion that these might predict which interventions are selected as optimal by the AI system for different participants. Understanding individual factors that impact intervention selection is potentially useful for future iterations of the artificial intelligence system, which could augment optimization with reward score data from groups of individuals identified as similar, especially early on when the system has relatively small amounts of data on a given individual. Future studies may also strengthen such a system by improving the usability and effectiveness of remote coaching (e.g., by having interventions that only occur once a week, simplifying texting protocol, and increasing tailoring of automated coaching by allowing refinement of goals by clinicians).

Even more fundamentally, there are many possible ways that the interventions could have been structured to allow for the variability needed for reinforcement learning to operate. For example, future studies could investigate the extent to which a reinforcement learning system can optimize treatments based on the training level (and thus cost) of the coach delivering the intervention (e.g., Ph.D.-level weight control expert or paraprofessional), the intervention content (e.g., informational versus supportive versus challenging; or, alternatively by selecting amongst different skill domains), prescription (e.g., related to diet or physical), intervention frequency, and intervention timing. Each of these intervention features could be expected to produce different levels of response by individual and also within individual, across time. Moreover, future reinforcement learning systems could experiment with complex combinations of intervention feature variation, and features could have more than the three levels chosen for this first proof-of-concept study. On the other hand, optimization will not be successful unless the variability of intervention is balanced against the number of times interventions are delivered to each participant. The current study was limited to 24 intervention occasions (12 weeks, two times per week), but this number could be greatly increased by increasing the frequency of interventions and extending the period of remote coaching. In particular, for longer-term iterations of this type of intervention, twice weekly contacts may become burdensome or unnecessary for some participants; thus, one “level” of intervention that could be integrated is a “no intervention” level.

Several limitations of the current study must be emphasized. Most importantly, this was a proof-of-concept, pilot trial with a relatively small sample size. As such, while the mean weight losses across conditions were similar in the remote coaching phase, the degree of confidence that results would generalize to the larger population is moderate, and the result should be replicated in a larger study. A related point is that the sample was primarily White, thus limiting generalizability. Participants were also asked to purchase materials for the study (e.g., scale and Fitbit) and strategies leveraging remote monitoring require an ability to proficiently use technology, which may further limit generalizability. Additionally, the remote coaching phase was only 3 months in duration, which provided relatively few periods for the system to learn about the effects of each intervention on each participant and to adapt to changing circumstances. The short duration of the trial leaves open the possibility that optimization would be even more effective in the long-term, but also that the benefits of AI fade over time. Several limitations were noted in connection with the intervention system. For example, the system depended, in part, on participants responding to coach text messages and phone calls in a limited window (i.e., during the time when coaches were available). Despite instructions to this effect and even pre-scheduling texts/calls to the minute, participants sometimes were not timely in their response. A later iteration of the system at a larger scale might widen coach availability windows and also let participants choose a specific time to engage. Increasing coach availability windows might also (a) help mitigate problems that were encountered in the individual optimization groups related to the fluctuations in total time a coach needs to commit on a given day and (b) address participants’ occasional need for longer phone calls to adequately overcome complex barriers to weight loss. Finally, while clinician time was limited by use of the optimization algorithms, clinicians reported contacting a large number of participants in a condensed period of time was challenging. The limit on efficiency and the systems needed to facilitate coaches’ effectiveness under high-volume conditions remain unknown.

In sum, the current study demonstrated that a remotely-delivered behavioral weight loss intervention could be successfully optimized for individual differences using an AI system that continuously scored, compared and selected varying interventions based on monitored digital physical activity, diet, tracking and weight data. This system should be further investigated given that it was able to deliver equivalent weight losses at a fraction of the cost.

Acknowledgements

This project was funded by a Drexel University DARE (Drexel Areas of Research Excellence) Award to Evan Forman. We would also like to acknowledge the assistance of Pedrag Krasnja whom we consulted during the planning of this study.

References

Abbeel P, & Ng AY (2004). Apprenticeship learning via inverse reinforcement learning. Paper presented at the Proceedings of the twenty-first international conference on Machine learning.

Anderson JW, Vichitbandra S, Qian W, & Kryscio RJ (1999). Long-term weight maintenance after an intensive weight-loss program. Journal of the American College of Nutrition, 18, 620–627. [PubMed: 10613414]

Association of Behavioral and Cognitive Therapies. (2018). Find a CBT Therapist. Retrieved May 23, 2018, from http://www.findcbt.org.

Auer P, Cesa-Bianchi N, & Fischer P (2002). Finite-time analysis of the multiarmed bandit problem. Machine Learning, 47, 235–256.

Aydin ME, & Öztemel E (2000). Dynamic job-shop scheduling using reinforcement learning agents. Robotics and Autonomous Systems, 33, 169–178.

Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, & Noakes M (2012). Features predicting weight loss in overweight or obese participants in a web-based intervention: Randomized trial. Journal of medical Internet research, 14, e173. [PubMed: 23234759]

Butryn ML, Forman EM, Lowe MR, Gorin AA, Zhang F, & Schaumberg K (2017). Efficacy of environmental and acceptance-based enhancements to behavioral weight loss treatment: The ENACT trial. Obesity, 25, 866–872. [PubMed: 28337847]

Butryn ML, Zhang F, Remmert JE, Roberts SR, & Forman EM (2018). Baseline Executive Functioning Predicts Weight Loss and Physical Activity Outcomes in A Lifestyle Modification Program. Paper presented at the 39th Annual Meeting and Scientific Sessions of the Society of Behavioral Medicine, New Orleans, LA.

Carels RA, Darby L, Cacciapaglia HM, Douglass OM, Harper J, Kaplar ME, et al. (2005). Applying a stepped-care approach to the treatment of obesity. Journal of Psychosomatic Research, 59, 375–383. [PubMed: 16310019]

Carels RA, Darby L, Cacciapaglia HM, Konrad K, Coit C, Harper J, et al. (2007). Using motivational interviewing as a supplement to obesity treatment: A stepped-care approach. Health Psychology, 26, 369. [PubMed: 17500624]

Carels RA, Hoffmann DA, Hinman N, Burmeister JM, Koball A, Ashrafioun L, et al. (2013). Step-down approach to behavioural weight loss treatment: A pilot of a randomised clinical trial. Psychology & Health, 28, 1121–1134. [PubMed: 23607537]

Carels RA, Selensky JC, Rossi J, Solar C, & Hlavka R (2017). A novel stepped-care approach to weight loss: The role of self-monitoring and health literacy in treatment outcomes. Eating Behaviors, 26, 76–82. [PubMed: 28189945]

Centers for Disease Control and Prevention (CDC). (2014). National Health and Nutrition Examination Survey Data. Hyattsville, MD Retrieved May 13, 2018, from https://www.cdc.gov/nchs/fastats/obesity-overweight.htm.

Diabetes Prevention Program Research Group. (2003). Costs associated with the primary prevention of type 2 diabetes mellitus in the diabetes prevention program. Diabetes Care, 26, 36–47. [PubMed: 12502656]

Forman EM, Butryn ML, Manasse SM, Crosby RD, Goldstein SP, Wyckoff EP, et al. (2016). Acceptance-based versus standard behavioral treatment for obesity: Results from the mind your health randomized controlled trial. Obesity, 24, 2050–2056. [PubMed: 27670400]

Jakicic JM, Tate DF, Lang W, Davis KK, Polzien K, Rickman AD, et al. (2012). Effect of a stepped-care intervention approach on weight loss in adults: A randomized clinical trial. JAMA, 307, 2617–2626. [PubMed: 22735431]

Jeffery RW, Epstein LH, Wilson GT, Drewnowski A, Stunkard AJ, & Wing RR (2000). Long-term maintenance of weight loss: Current status. Health Psychology, 19, 5. [PubMed: 10709944]

Joo N-S, & Kim B-T (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, 416–420. [PubMed: 18078554]

Korinek EV, Phatak SS, Martin CA, Freigoun MT, Rivera DE, Adams MA, et al. (2018). Adaptive step goals and rewards: A longitudinal growth model of daily steps for a smartphone-based walking intervention. Journal of Behavioral Medicine, 41, 74–86. [PubMed: 28918547]

Kramer FM, Jeffery RW, Forster JL, & Snell MK (1989). Long-term follow-up of behavioral treatment for obesity: Patterns of weight regain among men and women. International Journal of Obesity, 13, 123–136. [PubMed: 2663745]

Krukowski RA, Tilford JM, Harvey-Berino J, & West DS (2011). Comparing behavioral weight loss modalities: Incremental cost-effectiveness of an internet-based versus an in-person condition. Obesity, 19, 1629–1635. [PubMed: 21253001]

Kuderer M, Gulati S, & Burgard W (2015). Learning driving styles for autonomous vehicles from demonstration. Paper presented at the 2015 IEEE International Conference on Robotics and Automation (ICRA).

McGraa KLK (2010). The effects of persuasive motivational text messaging on adherence to diet and exercise programs across different personality traits. Santa Barbara: Fielding Graduate University.

Ng AY, Coates A, Diel M, Ganapathi V, Schulte J, Tse B, et al. (2006). Autonomous inverted helicopter flight via reinforcement learning In Ang MH & Khatib O (Eds.), Experimental Robotics IX (pp. 363–372). Berlin: Springer.

Ontanón S (2013). The combinatorial multi-armed bandit problem and its application to real-time strategy games. Paper presented at the Ninth Artificial Intelligence and Interactive Digital Entertainment Conference.

Ontanón S (2017). Combinatorial multi-armed bandits for real-time strategy games. Journal of Artificial Intelligence Research, 58, 665–702.

Paredes P, Gilad-Bachrach R, Czerwinski M, Roseway A, Rowan K, & Hernandez J (2014). PopTherapy: Coping with stress through pop-culture. Paper presented at the Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare.

Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. (2009). A text message–based intervention for weight loss: Randomized controlled trial. Journal of medical Internet research, 11, e1. [PubMed: 19141433]

Phatak SS, Freigoun MT, Martín CA, Rivera DE, Korinek EV, Adams MA, et al. (2018). Modeling individual differences: A case study of the application of system identification for personalizing a physical activity intervention. Journal of Biomedical Informatics, 79, 82–97. [PubMed: 29409750]

Piette JD, Farris KB, Newman S, An L, Sussman J, & Singh S (2014). The potential impact of intelligent systems for mobile health self-management support: Monte Carlo simulations of text message support for medication adherence. Annals of Behavioral Medicine, 49, 84–94.

Piette JD, Krein SL, Striplin D, Marinec N, Kerns RD, Farris KB, et al. (2016). Patient-centered pain care using artificial intelligence and mobile health tools: Protocol for a randomized study funded by the US Department of Veterans Affairs Health Services Research and Development Program. JMIR Research Protocols, 5, e53. [PubMed: 27056770]

Poirier J, Bennett WL, Jerome GJ, Shah NG, Lazo M, Yeh H-C, et al. (2016). Effectiveness of an activity tracker-and internet-based adaptive walking program for adults: A randomized controlled trial. Journal of Medical Internet Research, 18, e34. [PubMed: 26860434]

Rabbi M, Aung MH, Zhang M, & Choudhury T (2015). MyBehavior: Automatic personalized health feedback from user behaviors and preferences using smartphones. Paper presented at the Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529, 484–489. [PubMed: 26819042]

Snook KR, Hansen AR, Duke CH, Finch KC, Hackney AA, & Zhang J (2017). Change in percentages of adults with overweight or obesity trying to lose weight, 1988–2014. JAMA, 317, 971–973. [PubMed: 28267846]

Society of Clinical Psychology. (2018). Therapists Near Me. Retrieved May 23, 2018, from https://www.div12.org/therapistsearch/.

Sutton RS, & Barto AG (1998). Reinforcement learning: An introduction (Vol. 1). Cambridge: MIT press Cambridge.

Wadden TA, Butryn ML, & Wilson C (2007). Lifestyle modification for the management of obesity. Gastroenterology, 132, 2226–2238. [PubMed: 17498514]

Wadden TA, Sternberg J, Letizia K, Stunkard A, & Foster G (1989). Treatment of obesity by very low calorie diet, behavior therapy, and their combination: A five-year perspective. International Journal of Obesity, 13, 39–46. [PubMed: 2613427]

Webber KH, Tate DF, & Bowling JM (2008). A randomized comparison of two motivationally enhanced Internet behavioral weight loss programs. Behaviour Research and Therapy, 46, 1090–1095. [PubMed: 18675402]

Wilson GT (1994). Behavioral treatment of obesity: Thirty years and counting. Advances in Behaviour Research and Therapy, 16, 31–75.

Wilson GT, & Brownell KD (2002). Behavioral treatment for obesity In Fairburn CG & Brownell KD (Eds.), Eating disorders and obesity. New York: Guilford Press.

Yom-Tov E, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, & Hochberg I (2017). Encouraging physical activity in patients with diabetes: Intervention using a reinforcement learning system. Journal of Medical Internet Research, 19, e338. [PubMed: 29017988]

Zhou M, Fukuoka Y, Mintz Y, Goldberg K, Kaminsky P, Flowers E, et al. (2018). Evaluating machine learning-based automated personalized daily step goals delivered through a Mobile Phone App: Randomized controlled trial. JMIR mHealth and uHealth, 6, e28. [PubMed: 29371177]

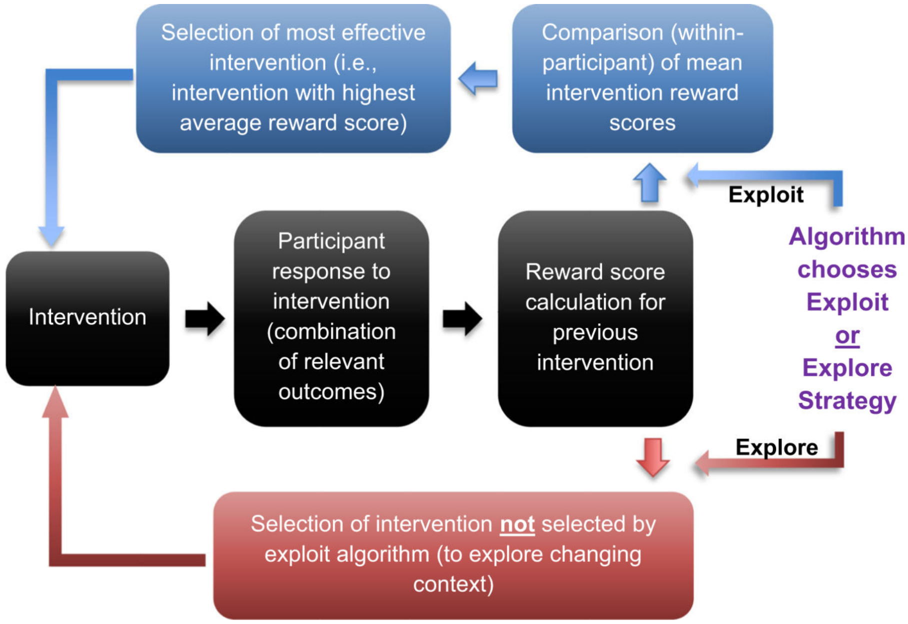

Fig. 1.

Reinforcement learning system which continuously monitors responses to interventions and repeatedly selects interventions, for every participant, based on the *exploit strategy* (use existing knowledge to select most effective intervention so far) or the *explore strategy* (select an alternate intervention to learn to allow for changing contexts)

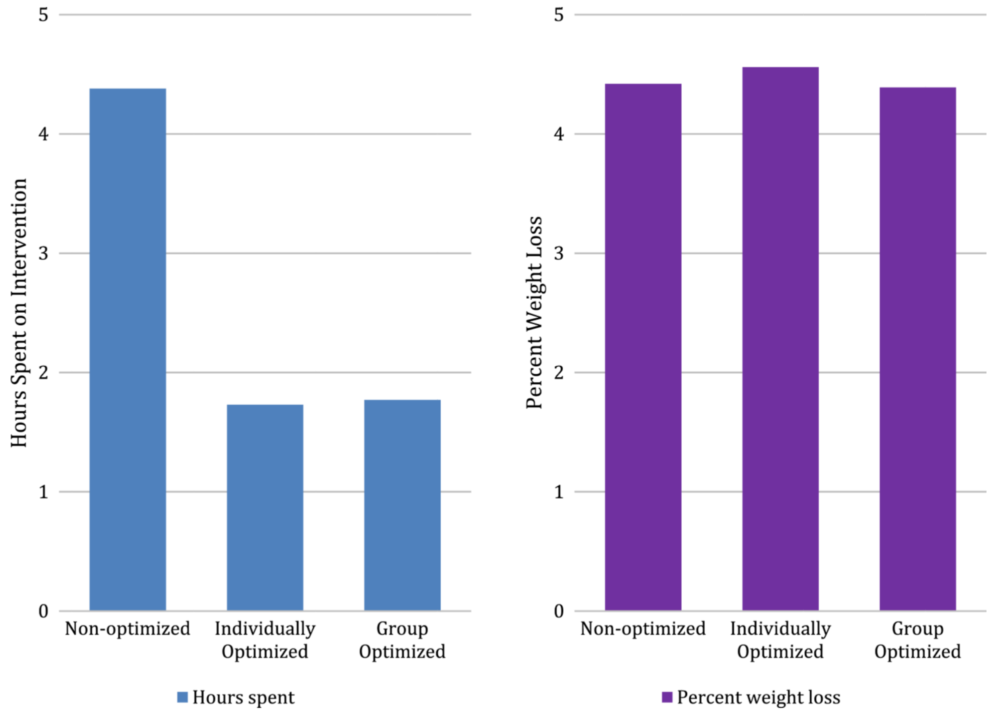

Fig. 2.
Time costs and weight losses by treatment condition

Table 1

Description of interventions potentially delivered

	Intervention time (in coach minutes)	Assigned to nonoptimized	Assigned to optimized	Description	Example
Automated text messages	0	No	Yes	Automated messages summarized the participant's progress (i.e., weight loss, self-weighing, food tracking, caloric intake, physical activity) in the past week. Messages then praised the participant and recommended a strategy to maintain progress in the following few days	*Hello. We will not be having a conversation today, but I wanted to touch base quickly on your progress. You lost 0.22 lbs in the last week. Over the last 4 days, you weighed yourself 3 days, tracked your food 2 days, met your calorie goal 1 day, and met your physical activity goal 2 days. Great work this week! Without tracking food intake, it's hard to know what to change to continue with weight loss. It's important that you start tracking regularly again. Weighing yourself daily is one of the major keys to long-term success. Set a reminder (like a phone alarm or a sticky note on your bathroom mirror). We'll check in again soon!*
Coach text exchange	3	No	Yes	Coaches had brief conversations with participants by text to check in on their progress, praise behavioral adherence, problem solve inadherence, and plan behavioral strategies for the upcoming days before the next coaching session	***Coach:*** *Hi ____. Good job the past few days weighing yourself each day and recording all of your food! I see you had a small gain. What do you think contributed to that?* ***Participant:*** *I guess not being active enough. I had an insane week taking care of my kids, since they were off school for the holiday.* ***Coach:*** *That must have been tough to juggle. One of the best ways to keep up your progress, though, is to get in some activity every day, even if it's a small amount! Do you have a plan to get back into your exercise routine soon?* ***Participant:*** *…*
Coach phone call	12	Yes	Yes	Content of the calls focused on checking in on participants' progress towards behavioral goals, reviewing strategies discussed during the 4-week workshop, praising behavioral adherence, problemsolving inadherence, and planning behavioral strategies for the upcoming days before the next coaching session	***Coach:*** *Hi ____. How have you been these past few days?* ***Participant:*** *I've been doing pretty well, thanks.* ***Coach:*** *Great job getting in lots of activity the past few days and weighing yourself each day. I took a look at your food records and noticed Wednesday was pretty high. What was going on that day?* ***Participant:*** *Yes, I thought about why I was so hungry that night, and I think it was because I had to spend more hours in the office than usual. I was really stressed and hungry by the time I got home. My schedule is back to normal now, though.* ***Coach:*** *That's great that you reflected on what went on since it was such a higher than normal day. It's good to figure out what all of the contributing factors are when we overeat. Was it challenging raising your exercise goal this week?* ***Participant:*** *…*

Table 2

Description of treatment conditions

Treatment condition	Description
Non-optimized	Phone coaching was assigned on each occasion for each participant
Individually optimized	Determined each occasion by highest reward score for each participant, adjusted by intervention cost
Group optimized	Determined each occasion by highest reward score sum for entire group that fits in a specified period of time (where time = number of participants * 5 min)

Table 3

Example of AI intervention selections for an individually-optimized

Intervention Period	Intervention	Step 1						Step 2			Step 3
		Days self-weighed	Days tracked food intake	Days met calorie goal	Days met physical activity goal	Weight change (lbs)	Reward score assigned to intervention	Average Costadjusted Autotext Reward Score (divided by 1)	Average Costadjusted Text Exchange Reward Score (divided by 1.1	Average Costadjusted Phone Call Reward Score (divided by 1.11)	Intervention selected for next period (selection strategy)
8/29–8/31	Text exchange	2	2	0	2	0.00	2.00	-	1.82	-	Text exchange (explore)
9/1–9/4	Text exchange	4	4	2	4	− 0.88	3.72	-	2.60	-	Phone call (explore)
9/5–9/7	Phone call	3	3	2	3	− 0.88	3.96	-	2.60	3.57	Autotext (explore)
9/8–9/11	Autotext	4	4	0	3	− 1.32	3.08	3.08	2.60	3.57	Phone call (explore)
9/12–9/14	Phone call	3	3	0	3	− 0.44	3.15	3.08	2.60	3.20	Autotext (explore)
9/15–9/18	Autotext	4	4	1	4	0.44	3.15	3.12	2.60	3.20	Phone call (exploit)
9/19–9/21	Phone call	3	3	0	3	2.21	2.19	3.12	2.60	2.79	Autotext (exploit)
9/22–9/25	Autotext	4	4	1	4	0.44	3.14	3.13	2.60	2.79	Autotext (exploit)

Values in this table are based on real participant data. Once an intervention was delivered, the following 3–4 days comprised that intervention period. During Step 1 (i.e., over the course of those 3–4 days), an individual's goals were tracked (self-weighing, tracking food, meeting calorie goal, and meeting physical activity goal) and weight change was calculated. For each day that each goal was met, a point was awarded. Weight change during this 3–4 day period was subtracted from that total (i.e., weight losses increased the reward score and weight gains decreased the reward score). The total reward score over that 3–4 day period was divided by the number of days in the block to generate the reward score for that intervention period. During Step 2, the average reward score for the most recent intervention block was averaged with the other reward scores achieved with that intervention, and the average reward score for that intervention was updated. In Step 3, the system considered the updated average reward scores for each intervention, and selected the intervention with the highest corresponding value

Table 4

Participant characteristics

	Non-optimized Mean (SD) or percent	Individually optimized Mean (SD) or percent	Group optimized Mean (SD) or percent	F or χ^2 (df)	p
N	12	20	20		
BMI at randomization	32.96 (6.48)	34.31 (5.24)	35.10 (5.96)	0.01 (2)	.99
Percent weight loss during Phase I	2.07 (1.66)	3.12 (2.06)	2.86 (1.85)	1.19 (2)	.31
Age	53.62 (13.96)	53.23 (12.97)	53.95 (14.11)	0.51 (2)	.60
Gender (% female)	66.7	60.0	75.0	1.03 (2)	.60
Race (%)					
White	72.7%	85.0%	80.8%	4.59 (6)	.60
Black	27.3%	10.0%	15.0%		
Asian	0.0%	5.0%	0.0%		
More than one race	0.0%	0.0%	5.0%		

Table 5

Characterization and comparison of interventions

	Non-optimized Mean (SD) (Range)	Individually optimized Mean (SD) (Range)	Group optimized Mean (SD) (Range)	F or χ^2	p (partial eta (df) sq or phi)
Percent of time algorithm selected automated texts	-	45.74 (14.92) (17.39–70.83)	44.61 (13.42) (25.00–75.00)	0.06 (1)	.80 (.002)
Percent of time algorithm selected text exchanges	-	28.23 (10.97) (12.50–52.17)	30.58 (11.97) (8.33–56.00)	0.42 (1)	.52 (.01)
Percent of time algorithm selected phone calls	-	26.04 (8.92) (13.64–50.00)	24.82 (9.77) (8.00–45.83)	0.17 (1)	.68 (.004)
Number of times optimal intervention (i.e., according to reward score) changed during the intervention	-	1.85 (1.87) (0–6)	2.75 (1.41) (0–5)	2.95 (1)	.09 (.07)
Daily behavioral adherence (number of behavioral goals met)	2.77 (0.60) (2.04–3.63)	2.73 (0.75) (0.73–3.70)	2.86 (0.49) (1.66–3.76)	0.23 (2)	.80 (.009)
Percent weight loss during remote treatment	4.42 (2.92) (0.27–10.23)	4.56 (3.22) (− 0.94–11.60)	4.39 (3.38) (− 0.46–11.54)	0.02 (2)	.98 (.001)
Percent of participants reaching at least 5% weight loss	58.3	65.0	70.0	0.45 (2)	.80 (.09)

Table 6

Mean number of behavioral goals met following each intervention type, by treatment condition

	Mean number (and SD) of behavioral goals met following...		
	Phone call	Text Exchange	Automated Text
Individually optimized	2.74 (0.76)	2.76 (0.78)	2.70 (0.76)
Group optimized	2.88 (0.52)	2.78 (0.45)	2.88 (0.54)