

Original Paper

Gamification of Behavior Change: Mathematical Principle and Proof-of-Concept Study

Falk Lieder^{1,2*}, PhD; Pin-Zhen Chen^{2*}, BSc; Mike Prentice², PhD; Victoria Amo², BSc, MSc; Mateo Tošić², BSc

¹Department of Psychology, University of California, Los Angeles, Los Angeles, CA, United States

²Max Planck Institute for Intelligent Systems, Tübingen, Germany

* these authors contributed equally

Corresponding Author:

Falk Lieder, PhD

Department of Psychology

University of California, Los Angeles

502 Portola Plaza, Los Angeles

Los Angeles, CA, 90095

United States

Phone: 1 424 259 5300

Email: falk.lieder@psych.ucla.edu

Abstract

Background: Many people want to build good habits to become healthier, live longer, or become happier but struggle to change their behavior. Gamification can make behavior change easier by awarding points for the desired behavior and deducting points for its omission.

Objective: In this study, we introduced a principled mathematical method for determining how many points should be awarded or deducted for the enactment or omission of the desired behavior, depending on when and how often the person has succeeded versus failed to enact it in the past. We called this approach *optimized gamification of behavior change*.

Methods: As a proof of concept, we designed a chatbot that applies our optimized gamification method to help people build healthy water-drinking habits. We evaluated the effectiveness of this gamified intervention in a 40-day field experiment with 1 experimental group (n=43) that used the chatbot with optimized gamification and 2 active control groups for which the chatbot's optimized gamification feature was disabled. For the first control group (n=48), all other features were available, including verbal feedback. The second control group (n=51) received no feedback or reminders. We measured the strength of all participants' water-drinking habits before, during, and after the intervention using the Self-Report Habit Index and by asking participants on how many days of the previous week they enacted the desired habit. In addition, all participants provided daily reports on whether they enacted their water-drinking intention that day.

Results: A Poisson regression analysis revealed that, during the intervention, users who received feedback based on optimized gamification enacted the desired behavior more often (mean 14.71, SD 6.57 times) than the active (mean 11.64, SD 6.38 times; $P < .001$; incidence rate ratio=0.80, 95% CI 0.71-0.91) or passive (mean 11.64, SD 5.43 times; $P = .001$; incidence rate ratio=0.78, 95% CI 0.69-0.89) control groups. The Self-Report Habit Index score significantly increased in all conditions ($P < .001$ in all cases) but did not differ between the experimental and control conditions ($P > .11$ in all cases). After the intervention, the experimental group performed the desired behavior as often as the 2 control groups ($P \geq .17$ in all cases).

Conclusions: Our findings suggest that optimized gamification can be used to make digital behavior change interventions more effective.

Trial Registration: Open Science Framework (OSF) H7JN8; <https://osf.io/h7jn8>

(*JMIR Serious Games* 2024;12:e43078) doi: [10.2196/43078](https://doi.org/10.2196/43078)

KEYWORDS

gamification; points; feedback; behavior change; habit formation; chatbot; digital interventions; mobile phone; artificial intelligence

Introduction

Background

Overview

People often struggle to change their behavior in ways that would benefit them in the long run. For instance, many people could improve their health and life expectancy by building healthy habits such as exercising [1], healthy eating [2], or drinking a glass of water before every meal [3,4]. People who want to adopt healthy habits because they know about their benefits nevertheless struggle to change their behavior accordingly.

Building a good habit is especially difficult when the benefits of the desired behavior cannot be felt immediately. One approach to alleviate this challenge is to create apps that encourage the desired behavior using incentives or immediate positive feedback [5-7] or discourage unwanted behavior using immediate negative feedback [8]. Doing so in a gameful way can be a promising approach to improving people's health behaviors [5-7,9-12]. Using game elements to improve people's behavior in the real world is known as gamification [13]. One of the most commonly used gamification methods is awarding people points for specific behaviors [14,15]. The awarded points are often used to provide feedback to the user, define levels, award badges, or create leaderboards.

Despite the widespread use of points, levels, badges, and leaderboards, there is currently no principled way to choose precisely how many points a person should be awarded and when. This is a problem as making those decisions based on intuition can lead to point systems that inadvertently incentivize counterproductive behaviors or undermine the users' motivation [16-18]. For instance, a recent study found that the point system of the popular gamified habit formation app Habitica is actively harmful [17].

To help practitioners master the challenge of designing effective point systems that reliably foster positive behavior change, we introduced a mathematical principle for computing the number of points a person should receive for engaging in the desired behavior depending on their history of either engaging or not engaging in this behavior and how many points they should lose when they fail to do so. We called this principle *optimized gamification of behavior change*. As a proof of concept, we applied optimized gamification to design a chatbot that helps people develop the healthy habit of drinking water before every meal [3,4]. Our chatbot awards points for the desired behavior and deducts points for its omission. Critically, the number of points that the user gains or loses is computed using optimized gamification. Our chatbot combines optimized gamification with three established principles of behavior change: it (1) guides the user to set an implementation intention, (2) reminds them of their good intentions, and (3) supports self-monitoring.

We evaluated our chatbot in a longitudinal field experiment and found that optimized gamification can make digital behavior change interventions more effective. Our findings provide a proof of concept for a very general and principled approach to improving human behavior in the real world. In addition to this

theoretical contribution, we introduced a chatbot for helping people develop a specific healthy habit (water drinking): the Good Habit Bot. This chatbot can easily be adapted to other health behaviors, including more critical health behaviors such as exercising, healthy eating, and other good habits that people want to establish.

The plan for this paper is as follows. We first introduce relevant background information about behavior change and gamification. We then present our theory of optimized gamification for behavior change. Next, we present the chatbot we designed as a proof of principle. After that, we present the methods and results of our field experiment and discuss its findings and implications.

Behavior Change Goals Versus Automatic Behavior

Human behavior is controlled by a combination of goal-directed decision-making (eg, *I will buy a gym membership because I want to lose weight*) and more automatic reactions to certain stimuli (eg, always stopping by the gym on the way home from work) [19]. Goal-directed decision-making derives choices from the outcomes that people value (eg, health or money) or want to avoid (eg, pain) according to their mental models of how those outcomes can be obtained. In contrast, automatic reactions do not consider the behavior's likely consequences in the current situation.

Obstacles to Behavior Change: Automaticity and Temporal Discounting

Most of our behavior is not primarily controlled by goal-directed decision-making but determined by people's automatic reactions. Therefore, automatic behavioral responses can interfere with people's ability to act in accordance with their behavior change goals. This is the proverbial conflict between bad habits and good intentions. People can inhibit and override their automatic behavioral responses (*bad habits*), but their capacity to do so is limited [20]. Therefore, the automaticity of human behavior is a crucial obstacle to intentional behavior change.

A second obstacle to successful behavior change is that the mechanisms of goal-directed decision-making are demonstrably biased in favor of immediate outcomes [21]. People give too much weight to their decision's immediate consequences and too little weight to its long-term consequences. This phenomenon is known as present bias [22]. It has been proposed that present bias occurs because the brain discounts delayed benefits as if they become less valuable the later they occur [23]. This phenomenon, which is known as temporal discounting, is well established in research on economic decisions and animal behavior.

Moreover, according to temporal motivation theory [24], temporal discounting is one of the main reasons why people fail to enact good intentions. Such failures of self-regulation are a critical obstacle to health behavior change [25,26]. Consistent with this explanation, people who discount delayed outcomes more strongly are likelier to engage in unhealthy behaviors and experience poor health [26].

Reinforcement Learning as a Mechanism of Behavior Change

As automatic responses are powerful drivers of human behavior, successful behavior change typically involves translating behavior change goals into automatic behavioral responses [27].

Automatic behavioral responses, including exercise habits, can be acquired through learning from experience. Model-free reinforcement learning is a well-established mechanism of learning automatic behavioral responses from experience [19]. This mechanism increases or decreases a person's propensity to engage in a specific behavior in a particular situation (eg, going to the gym after work) according to whether they experience the behavior's overall consequences as positive or negative. The vast literature on operant conditioning in animals and humans underscores that learning from reward and punishment is a powerful mechanism of behavior change [28,29]. Another complementary learning mechanism involves strengthening habits through mere repetition [30].

Supporting Behavior Change Through Incentives and Reinforcement

The literature surveyed previously demonstrates that goal-directed decision-making and automatic behavioral responses are responsive to rewards and punishments. Goal-directed decision-making is sensitive to anticipated future rewards, and automatic behavioral responses are shaped by the rewards or punishments that those behaviors have generated in the past. These effects can be leveraged to support behavior change. To foster behavior change via goal-directed decision-making, behavior change interventions can create and announce incentives for engaging in the desired behavior. To foster behavior change via reinforcement learning, behavior change interventions can reinforce the desired behavior with rewards or positive feedback.

A highly effective behavior change intervention that leverages both effects is contingency management [31]. Contingency management incentivizes a desired behavior change and rewards people when they enact it. Voucher-based reinforcement therapy for treating addiction is a highly successful example of contingency management [32]. This behavior change intervention awards the patient a voucher every time they submit a negative drug test. More recently, it has also been applied to foster other types of behavior change, including physical exercise [7,33] and treatment attendance [34]. Contingency management appears to be more effective when the desired behavior is reinforced more promptly, more frequently, and with rewards that are larger or increase throughout the intervention [35-37].

Digital Behavior Change Interventions

Developing digital behavior change interventions is a young and booming field [38]. Mobile apps have shown potential for fostering positive behavior change in domains such as physical exercise and healthy eating [38,39]. However, the average effect size of such interventions is still relatively small [40,41]. Most behavior change interventions are not derived from any theory, model, or framework [38]. Therefore, we suspect that there is still room for improvement and that at least some of this

potential can be realized by adopting a more theory-driven approach.

Goal setting and self-monitoring are the most commonly used behavior change techniques [38]. A meta-analysis of studies on digital interventions for promoting physical exercise found that these 2 techniques are also the most effective ingredients of current digital behavior change interventions [41]. Goal setting entails guiding people to articulate their intent to perform certain behaviors in certain situations (eg, drinking a glass of water before every meal). Supporting self-monitoring often takes the form of helping people check or record whether they have enacted those intentions or track related outcomes (eg, their weight). Goal setting is especially effective when people formulate simple plans that specify the intended behavior and the situation in which they want to perform it as concretely as possible [42,43]. This approach is known as *implementation intentions*. Moreover, reminding people of their intentions via SMS text messages [44] and presenting them with positive reinforcement when they enact their intentions [5,7,33] have been found to be highly effective in promoting physical exercise.

Gamification

A recent meta-analysis found that approximately 1 in 5 digital behavior change interventions are designed within the gamification framework [38]. Gamification entails applying principles from game design and game elements, such as storytelling and rules for earning points and winning the game, to address real-world problems [13]. The basic idea is to motivate people to do things that benefit them or others, such as exercising and studying, in a gameful way. Previous research has found that gamification can improve people's behavior, achieve desired outcomes, and improve people's subjective experiences [45].

Gamification is already widely used in designing digital behavior change interventions [5], and previous studies have suggested that it can improve people's health behaviors [7,9-12,33]. One gamification strategy that is effective in digital behavior change interventions is awarding the user points as positive feedback for the desired behavior [5,7,33]. Such extrinsic incentives can increase the frequency of the targeted behaviors without affecting people's intrinsic motivation [46,47].

However, when gamification is not correctly designed, it can backfire and have adverse effects [16,17,48]. This has also been observed in the behavior change literature [12,49] and in gamified habit formation apps [17]. Getting the incentives exactly right can be crucial as points, levels, badges, and leaderboards do not foster the user's intrinsic motivation [46,47] and might sometimes even undermine it when they are not embedded in a compelling narrative [18]. Motivated by these problems, many authors have called for a more theory-driven approach to gamification in general [50] and gamifying digital behavior change interventions in particular [12].

Optimized Gamification

Previous work has investigated how many points should be awarded for which behavior to maximally benefit the user in the context of to-do list apps that help the user achieve their

own goals [51,52]. Building on temporal motivation theory [24], this work assumed that people's motivation is insufficiently sensitive to long-term benefits such as good health in old age and overly sensitive to immediate costs (eg, the effort of exercising) and short-term pleasure (eg, from receiving immediate positive feedback). To help people overcome the resulting motivational problems (eg, procrastination) [53], Lieder et al [51] developed a mathematical theory for designing point systems that provide immediate positive feedback for activities that are beneficial in the long run and immediate negative feedback for activities that are not. The basic idea is to align each action's immediate and long-term consequences. The action that is best in the long run should be made most appealing in the short run, and actions with undesirable long-term consequences should be made unappealing in the short run.

Therefore, optimized gamification strives to incentivize each of the available actions through a number of points proportional to how much that action increases or decreases the sum total of future happiness. This idea is implemented by modeling the activities to be incentivized as steps that lead toward a valuable goal or away from it. Actions that lead toward the goal increase the time the person will spend in the more valuable state in which the goal has been achieved. In contrast, actions that lead away from the goal decrease the time the person will spend in the more valuable state in which the goal has been achieved and increase the effort required to achieve it afterward. On the basis of this mathematical model, dynamic programming and reinforcement learning methods can estimate how much a given action improves or worsens the person's situation. These estimates are then translated into incentives that encourage good choices and discourage bad ones. The resulting point values are optimal in that they would enable even a purely myopic decision maker who only cares about immediate outcomes to choose the actions that are best for them in the long run [51].

Although optimized gamification construes points as incentives and uses mathematical and computational methods from the field of reinforcement learning, using it does not constitute a commitment to behaviorism and is fully compatible with cognitive theories of motivation [54].

Optimized gamification has been used to encourage users to tackle the tasks on their to-do lists [32] and encourage students to select the most valuable learning activities [55,56]. Optimized gamification has also been applied to give people feedback on how they think about what to do [57] and on whether they succeeded in staying focused on a chosen task or got distracted [58]. However, to date, this approach has never been applied to support habit formation.

Objectives

The first goal of this study was to introduce a principled method for computing feedback on the enactment or omission of the desired behavior and experimentally test whether it can be used to enhance digital behavior change interventions. The second goal of this study was to introduce a chatbot that uses this method to help people develop healthy water-drinking habits and evaluate it in a longitudinal field experiment.

Methods

Optimized Gamification of Behavior Change

Overview

We conceptualized behavior change as a special case of repeatedly choosing and learning when to do what. As reviewed in the *Background* section, optimized gamification can encourage desired behaviors and accelerate learning [57]. To apply this method to promote the desired behavior and accelerate the formation of healthy habits, we first have to model habit formation as a Markov decision process (MDP) [59].

Modeling Habit Formation

An MDP is a scenario in which an agent faces a series of choices. Each choice (a) has 2 effects. First, it yields an immediate reward (r) that may be positive, negative, or zero. Second, it may change the state (s) the agent finds itself in. In an MDP, the agent's goal is to maximize the sum of the rewards it accumulates from its first decision to its last one.

We model behavior change problems as a straightforward MDP, in which a person repeatedly chooses between 2 possible actions when they find themselves in a particular situation: enacting the desired behavior ($a=1$) or not enacting it ($a=0$). Our model assumes that a given behavior change intervention aims to turn the desired behavior into a habit. Therefore, we define the state as the strength of the person's healthy habit, measured using a single number, s_{habit} , which can range from 0 to 1. Following standard habit formation models [30], we assume that enacting the habit increases its strength from

$$s_{\text{habit}} \text{ to } s_{\text{habit}} + \alpha \times (1 - s_{\text{habit}}),$$

where α is a free parameter that describes how quickly habits form. Conversely, our model assumes that failure to enact the desired behavior decreases the strength of the habit to $s_{\text{habit}} \times (1 - \alpha)$. We assume that the habit has been cultivated when its strength exceeds some threshold θ (eg, $\theta=0.9$) and model the health benefits conferred by achieving this goal as a reward (r_{goal}) that is attained when the habit strength crosses this threshold. Enacting the desired behavior is assumed to incur a cost that decreases with the strength of the habit ($r[s_{\text{habit}}, 1]=-[1 - s_{\text{habit}}]$), whereas not performing the behavior is assumed to be effortless ($r[s_{\text{habit}}, 0]=0$).

Computing Optimal Feedback

The basic idea of optimized gamification is to reward each action using a number of points that are proportional to its long-term benefits. These long-term benefits are measured via the decrease in future costs and the increase in future rewards brought about by transitioning to a state in which the habit is stronger. In situations in which the benefits of developing the good habit outweigh its costs, the value of having a habit of strength s_{habit} and then following through with the process of building the habit is as follows:

$$V^*(s_{\text{habit}}) = r_{\text{goal}} - (1 - s_{\text{habit}}) \cdot \sum_{i=1}^{n(s_{\text{habit}};\theta)} (1 - \alpha)^{i-1}$$

where the number $n(s_{\text{habit}}; \theta)$ specifies how often the behavior must be enacted until the habit strength reaches its target value θ . Therefore, for someone who will follow through with building the habit, the long-term benefits of enacting the habit one more time when its current strength is s_{habit} are

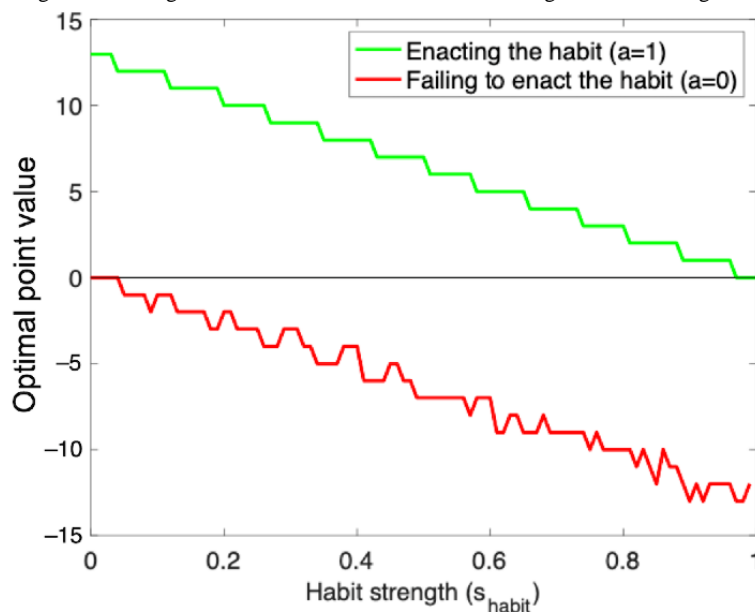
$$f(s_{\text{habit}}, 1) = V^*(s_{\text{habit}} + \alpha \times [1 - s_{\text{habit}}]) - V^*(s_{\text{habit}}) = 1 - s_{\text{habit}}.$$

Conversely, the long-term costs of failing to enact the desired behavior in the situation in which it is supposed to become a habit are

$$f(s_{\text{habit}}, 0) = V^*(s_{\text{habit}} \times [1 - \alpha]) - V^*(s_{\text{habit}}).$$

Please note that, even though we are talking about a situation in which it is rational for people to build the habit, this does not mean that we assume people to be rational. For our method, it does not matter why people follow through with building the habit. In fact, we assume that some people will follow through with building the habit only because they are (irrationally strongly) motivated by the immediate rewards conferred by feedback.

Figure 1. Point values for enacting versus failing to enact the habit at different habit strengths for a learning rate of $\alpha=0.1$ and a threshold of $\theta=0.9$.



As illustrated in Figure 1, the number of points for enacting the desired behavior is largest when the habit is weakest and gradually decreases toward 0 as the habit strengthens. This is intuitive as performing the desired behavior becomes easier the more often one has already performed it. Conversely, the number of points that should be deducted when the user fails to enact the habit is smallest when the strength of the habit is 0 and then increases as the habit becomes stronger. This is an intuitive consequence of our model's assumption that failing to enact the desired behavior in the specified situation weakens the habit strength by approximately 10%. The stronger the habit, the more is lost when its strength is reduced by 10%. The number of points awarded for enacting the desired behavior is a monotonically decreasing function of the habit strength. In contrast, the point value for failing to enact the desired behavior changes more erratically. This is because the number of steps required to reach the desired habit strength changes abruptly

The lowest possible negative value of f is $f(\theta, 0)$, and the largest possible positive value is $f(0, 1)$. Although the exact values depend on the model parameters, they are typically approximately -1 and 1 , respectively. Therefore, to transform those values into points, it is desirable to scale them by the desired maximum point value (M) that the application should award to the user and then round the scaled values to the nearest integer. This yields the following equation for the number of points that the application should award when a user reports that they have enacted their intention ($a=1$) or have not enacted their intention ($a=0$):

$$\text{points}(s_{\text{habit}}, a) = \text{round}(M \times f(s_{\text{habit}}, a)) \text{ (Equation 1).}$$

Figure 1 illustrates the point values as a function of the user's habit strength in an example application with a learning rate of $\alpha=0.1$, a target habit strength of $\theta=0.9$, and a maximal point value of $M=13$. Multimedia Appendix 1 provides more details on our mathematical model and the resulting optimal point system.

with the current habit strength. For instance, failing to enact the behavior at a habit strength of 0.09 increases the number of times the behavior needs to be enacted to achieve the desired habit strength from 21 to 22 times. However, if the habit strength is 0.08 or 0.10, the number of times the desired behavior has to be enacted remains 22 and 21 times, respectively. Individual users rarely experience such irregular changes as the change in their habit strength typically skips across those small areas in which the point value changes nonmonotonically. Moreover, our simulations suggest that the penalty for failing to enact the desired behavior can be approximated using a linear function with the same slope as the number of points for performing the desired behavior.

Application to Supporting Positive Behavior Change

The optimized gamification method described previously can be applied to help people form good habits. The equations for

computing the number of points are easy to implement within digital behavior change interventions such as chatbots and habit trackers. All that is needed is to ask the user which habit they want to develop and estimate its initial strength, set the learning rate parameter and the habit's target strength to reasonable values (eg, $\alpha=0.1$ and $\theta=0.9$), and record when the user did versus did not act in accordance with the desired habit. To define the desired habit, the user has to specify the desired behavior and the situation in which they want to perform it. The user's initial habit strength can be estimated through the desired behavior's relative frequency in that situation in the previous weeks (eg, at 2 of the previous 7 lunches). The habit strength can then be initialized using that proportion. Alternatively, when it makes sense to assume that the user wants to build an entirely new habit, the strength can be initialized as 0. Then, whenever the user reports having or not having enacted their intention, the optimized gamification equation can be applied to compute how many points the user should gain or lose. Whenever the user reports having performed the desired behavior, the estimate of the habit strength should be increased to

$$s_{\text{habit}} + \alpha \times (1 - s_{\text{habit}}),$$

and whenever the user reports having missed an opportunity to enact their intention, the estimate of their habit strength decreases to $s_{\text{habit}} \times (1 - \alpha)$. Then, the same procedure repeats when the user reports on their next opportunity to enact the behavior.

This approach can be used to support many different types of positive behavior change. It can help people build good habits in areas such as exercise, sleep, taking medication, nutrition, work, chores, and leisure activities. It can be used in interventions focusing on specific habits and, in general, in habit formation tools that let users choose any habit they want to develop. Another possible application is helping people overcome bad habits (eg, smoking cessation) [60]. It can support applications that run on practically any device, from smartwatches and wristbands to mobile phones, desktop applications, web applications, and smart glasses.

Proof of Concept: A Chatbot for Building a Healthy Water-Drinking Habit

As a proof of concept for the application of optimized gamification to support behavior change, we implemented this idea as a Telegram (Telegram FZ LLC) chatbot called the *Good Habit Bot*. This chatbot helps the user develop a healthy water-drinking habit by combining 4 behavior change techniques: goal setting, reminders, support for self-monitoring, and feedback. Concretely, the Good Habit Bot guides the user to formulate an implementation intention that links a specific desired behavior to a concrete daily event, reminds the user of their intention on a daily basis, checks in with them on whether they followed through on their intention every day, and then gives them positive or negative feedback depending on whether they did or did not follow through on their intention.

When the user starts their first conversation with the Good Habit Bot, the chatbot introduces itself and says that its purpose is to

help the user form a healthy water-drinking habit. The Good Habit Bot then asks the user to choose which of 8 concrete, recurring moments in their day they want to use as the trigger for their water-drinking habit (eg, *when my wake-up alarm rings* or *when I have the first bite of my lunch*; for the complete list, see [Multimedia Appendix 2](#)). Next, the Good Habit Bot asks the user how much water they want to drink in that situation (eg, 1 glass or 0.5 glasses) and how often they did so in the previous 7 days. The chatbot then uses the number of days n of the previous week in which the user acted in accordance with the habit (eg, $n=2$) to initialize their habit strength by $n/7$ (eg, $s_{\text{habit}} = 2/7$). Then, in the evening of the first day (ie, at 9 PM), the chatbot reminds the user of their intention to drink a specific amount of water at a particular moment of the following day (eg, *Remember your intention: When I have the first bite of my lunch, I will drink 1.5 glasses of water*). Then, sometime after the moment in which the user wanted to enact their intention, the chatbot asks them whether they did so (ie, *Did you accomplish your goal today to drink 1.5 glasses of water*). If the user affirms that they followed through on their intention, the chatbot gives them positive feedback ([Figure 2](#)). This feedback comprises a congratulatory message (eg, *That's wonderful!*) whose text alternates among 5 possible phrases ([Multimedia Appendix 2](#)) and a second message that awards the user the number of points computed by our optimized gamification method (eg, *I am glad to grant you 5 points for keeping a good habit! Your total score is 49 points*). In contrast, if the user responds that they missed their chance to enact their water-drinking intention, the Good Habit Bot tells them *Okay. Keep going tomorrow!* and informs them how many points they lost and how many they have left ([Multimedia Appendix 2](#)). Afterward, the chatbot updates the user's habit strength. Later that day, the Good Habit Bot reminds the user of their intention for the next day, and then the cycle repeats.

Critically, the chatbot computes how many points to award or take away from the user according to the optimized gamification method described previously (equation 1). One can read the number of points the chatbot we used in our experiment awarded in different situations in [Figure 1](#) as it used the same set of parameters (ie, $\alpha=0.1$; $\theta=0.9$; $M=13$). For instance, if a user who reported having performed the desired behavior twice in the previous week enacted their intention on the first day, they earned 9 points, and their habit strength increased from

$$\frac{2}{7} \approx 0.29 \text{ to } \frac{2}{7} + 0.1 \cdot \left(1 - \frac{2}{7}\right) \approx 0.36$$

Conversely, if they failed to enact their intention, they would lose 4 points, and their habit strength would decrease to

$$0.9 \cdot \frac{2}{7} \approx 0.26$$

The Good Habit Bot is freely available on the Telegram messenger app, where it can be found by searching for its alias, [@learn_good_habits_bot](#).

Figure 2. Screenshot of the feedback method with point values computed by our optimized gamification method.



Study Design

To evaluate the effectiveness of our intervention and assess the relative contributions of reminders versus optimized gamification, we ran a longitudinal between-subject experiment with 1 experimental condition with optimized gamification (*optimized gamification condition*), 1 *baseline condition*, and 1 *control condition with reminders and feedback* (Table 1). In the *optimized gamification condition*, the chatbot delivered all 4 techniques described previously, including feedback messages based on optimized gamification (Figure 1). Participants in the other 2 conditions could not gain or lose any points (no optimized gamification). They differed in whether they received feedback messages for enacting versus failing to enact the intended behavior. In the *control condition with reminders and feedback*, participants received a positive feedback message

when they reported having enacted their intention (eg, *That's wonderful!*) and a more neutral message when they reported not having done so (eg, *Okay. Keep going tomorrow!*). To create the *baseline condition*, we replaced the first control group's positive and negative feedback message with a neutral message (*OK*) and removed the daily reminders. Participants completed the self-report measures of habit strength described in the following sections before the intervention (pretest time point), immediately after the intervention (posttest time point), and approximately 3 weeks later (follow-up). Moreover, participants reported how often they engaged in the desired behavior the week before the intervention (pretest time point) and the week after the intervention (posttest time point). Finally, participants also completed daily reports of whether they enacted their intention on each day of the intervention.

Table 1. Experimental conditions.

Experimental condition	Implementation intentions	Support for self-monitoring	Reminders	Feedback
Optimized gamification condition	Yes	Yes	Yes	Optimized gamification and positive vs negative text
Control condition with reminders and feedback	Yes	Yes	Yes	Positive vs negative text
Baseline condition	Yes	Yes	No	None

Ethical Consideration

This experiment was conducted according to study protocol 401/2020BO2 approved by the Independent Ethics Commission at the Faculty of Medicine of the University of Tübingen. All data was collected and handled according to the General Data Protection Regulation of the European Union. All data has been de-identified.

Recruitment and Reimbursement

We recruited 132 participants (n=41, 31.1% for the baseline condition; n=43, 32.6% for the optimized gamification condition; n=48, 36.4% for the control condition; n=93, 70.5% female) on the web-based research platform Prolific. Based on considerations about the cost of the study, the sample size was

determined a priori so that we would retain 3x40=120 participants after an anticipated 10% of participants dropped out of the study. The requirements for participation were being a native English speaker, not having interacted with our chatbot before, and having previously completed at least 10 prolific assignments with an approval rate of at least 95%. Moreover, participants had to be aged ≥ 18 years. The average age of the participants was 31.5 (SD 9.9; range 19-79) years.

The study description informed participants about the study's duration, activities, time commitment, and pay. Participants were paid £1.95 (US \$2.45) for completing the onboarding survey. We informed them that the base pay for completing the remainder of the 40-day study would be £7.80 (US \$9.81) and that they could earn an additional bonus of £8 (US \$10.06). The

description strongly recommended that only people who were already using Telegram on their smartphones should participate. Moreover, the study description informed potential participants about the potential health benefits of regular water drinking. Participants were then shown the consent form. Upon providing informed consent, participants who already had the Telegram app clicked on a link that started a conversation with the Telegram chatbot for their corresponding experimental condition. Participants who had not installed the Telegram app yet were directed to download it first.

At the end of the 40-day study, the chatbot directed participants to a second Prolific HIT where they received the announced payments contingent on their sustained active participation. All participants who completed the pretest, posttest, and follow-up measures and continued to report their daily intention enactment after the 10th day received a second payment of £15.80 (US \$19.86). Participants who did not meet these criteria did not receive the second payment.

Outcome Measures and Procedure

Outcome Measures

We measured the outcome variables described in the following sections.

Dropout

We measured whether a participant dropped out of our study using a binary variable indicating whether the participant stopped responding to all daily reports at least 3 days before the end of the study.

Engagement

We measured each participant's *engagement* with our digital intervention based on the number of days on which they interacted with the Good Habit Bot.

Self-Report Habit Index

The Self-Report Habit Index (SRHI) [61] is a 12-item self-report measure of habit strength on a 7-point Likert scale. It comprises 3 subscales measuring the behavior's *history of repetition*, its *automaticity*, and the extent to which it is part of the person's *identity*. In this study, we administered the first 2 subscales. The SRHI has been found to be a 1D construct. Therefore, we averaged the scores of all items. The SRHI has been found to have high validity and very high reliability (Cronbach α of approximately 0.90; test-retest reliability: $r=0.91$).

Daily Intention Enactment

To measure how often each participant enacted their water-drinking intention during the intervention, we asked them the following question—*Did you accomplish your goal to drink 1 glass of water?*—on each day of the intervention. The question was asked between 30 minutes and 2.5 hours after the time at which the participant intended to drink water. Participants responded by selecting between the responses *Yes, I did!* and *No, I didn't*. After the study, we calculated each participant's *daily intention enactment* score by counting on how many of the 21 days of the intervention period they enacted their

intention. Therefore, the daily intention enactment score could range from 0 to 21.

Retrospective Intention Enactment

To measure how regularly participants engaged in the intended behavior (eg, drinking a glass of water before lunch) in the weeks before and after the intervention, we asked them to answer the following question—*On how many days of the previous week did you keep the habit of drinking water?*—by selecting one of the answer choices (0, 1, 2, 3, 4, 5, or 6 days). We referred to the resulting number of days as the *retroactive intention enactment* score.

Procedure

We created 3 separate Prolific HITs for each of the 3 conditions of the experiment (Table 1), and each person was allowed to participate in at most one of these HITs. The experiment ran from November 11, 2021 to December 19, 2021. As illustrated in Table 2, the experiment was divided into 3 phases: the preintervention period (day 0), intervention period (days 1-21), and postintervention period (days 22-40). In the preintervention period, participants provided informed consent, completed the onboarding process, and completed the pretest. We blinded participants to the experimental manipulation by giving all participants the same information about the chatbot they were interacting with and the anticipated benefits of interacting with it. During onboarding, participants were directed to start the first conversation with the Good Habit Bot in the Telegram app on their mobile phones. In this initial conversation, the chatbot asked the participants to select a concrete daily situation in which they wanted to drink water and how much water they wanted to drink, as described previously. The pretest comprised 2 self-report measures: the *SRHI* and the *retrospective intention enactment* measure for the week before the study.

During the intervention period, each participant interacted with 1 of the 3 versions of our chatbot according to the condition they were in (Table 1). On each day of the intervention period, all 3 groups reported whether they successfully enacted their intention to drink water in the specific situation they had selected (*daily intention enactment*). At the end of the intervention period, all 3 groups completed the *SRHI* for the second time (posttest time point).

The postintervention period started with a 1-week break during which the chatbot did not communicate with the participants. Then, on day 28 (follow-up 1) and day 35 (follow-up 2), the chatbot asked participants from all 3 groups to report on how many days of the previous week (ie, the first and second week of the postintervention period, respectively) they had acted in accordance with the desired habit (*retroactive intention enactment*). Finally, on day 40, the chatbot asked all participants to complete the *SRHI* questionnaire for the third time (follow-up). Participants received 3 email reminders to resume interacting with the chatbot on the day of the first follow-up survey (day 28), the day of the second follow-up survey (day 35), and the day of the final follow-up survey (day 40).

Table 2. Experimental procedure.

Experimental phase and day	Optimized gamification condition	Control condition with reminders and feedback	Baseline condition
Before the intervention			
Day 0	<ul style="list-style-type: none"> Onboarding Retrospective intention enactment SRHI^a 	<ul style="list-style-type: none"> Onboarding Retrospective intention enactment SRHI 	<ul style="list-style-type: none"> Onboarding Retrospective intention enactment SRHI
Intervention			
Days 1-21	<ul style="list-style-type: none"> Reminder, report, and feedback (optimal points+text) Daily intention enactment 	<ul style="list-style-type: none"> Reminder, report, and feedback (text only) Daily intention enactment 	<ul style="list-style-type: none"> Daily intention enactment
Day 21	<ul style="list-style-type: none"> SRHI 	<ul style="list-style-type: none"> SRHI 	<ul style="list-style-type: none"> SRHI
After the intervention			
Days 22-27	<ul style="list-style-type: none"> No reports 	<ul style="list-style-type: none"> No reports 	<ul style="list-style-type: none"> No reports
Day 28	<ul style="list-style-type: none"> Retrospective intention enactment 	<ul style="list-style-type: none"> Retrospective intention enactment 	<ul style="list-style-type: none"> Retrospective intention enactment
Days 29-34	<ul style="list-style-type: none"> No reports 	<ul style="list-style-type: none"> No reports 	<ul style="list-style-type: none"> No reports
Day 35	<ul style="list-style-type: none"> Retrospective intention enactment 	<ul style="list-style-type: none"> Retrospective intention enactment 	<ul style="list-style-type: none"> Retrospective intention enactment
Days 36-39	<ul style="list-style-type: none"> No reports 	<ul style="list-style-type: none"> No reports 	<ul style="list-style-type: none"> No reports
Day 40	<ul style="list-style-type: none"> SRHI 	<ul style="list-style-type: none"> SRHI 	<ul style="list-style-type: none"> SRHI

^aSRHI: Self-Report Habit Index.

Data Analysis

The hypotheses and statistical analysis plan were preregistered on the internet [62]. Participants who completed 0 daily water consumption reports were excluded from all analyses apart from the dropout analysis. Other than that, all analyses were conducted on all available data from all participants who completed at least one daily water consumption report. We retained 37/41 participants from the baseline condition, 39/43 participants from the control condition with reminders and feedback, and 42/48 participants from the optimized gamification condition. All comparisons between conditions were based on the originally assigned groups. We used Poisson regression analyses for binary outcome variables. For continuous outcome variables, we used linear multilevel modeling.

Results

There was no indication of failure of random assignment for initial habit strength, automaticity, or history of repetition (pairwise $P > .14$ in all cases).

Dropout

As an initial step, we examined whether there was a differential dropout among the 3 conditions using a chi-square test of independence. There was no effect of condition on dropout ($\chi^2_2=0.7$; Cramer $V=0.07$; $P=.72$).

Engagement

The *engagement* variable was entered into a Poisson regression model with 2 dummy variables for the effects of optimized gamification and feedback and reminders. The *automaticity* and *history of repetition* scores from the SRHI before the intervention and the preintervention *retrospective intention enactment* score were entered as control variables. As shown in Table 3, optimized gamification did not increase engagement compared with the *baseline condition*. However, it appears that being in the control condition with reminders and feedback without optimal points reduced *engagement* compared with the *optimized gamification* condition (Table 3) and the *baseline* condition (incidence rate ratio=0.84, 95% CI 0.75-0.94; $P=.002$).

Table 3. Predicting the number of days on which participants engaged with the app (engagement) from their condition and time-1 habit-related control variables (N=126)^a.

Predictor	Incidence rate ratio (95% CI)	P value
Intercept	18.88 (16.61-21.43)	<.001
Baseline vs optimized gamification	1.02 (0.91-1.14)	.72
Reminders and feedback vs optimized gamification	0.86 (0.76-0.96)	.008
Preintervention habit strength	1.03 (1.00-1.07)	.07
History of repetition	0.85 (0.79-0.92)	<.001
Automaticity	1.08 (1.01-1.16)	.03

^aNagelkerke $R^2=0.33$. The treatment group with optimal points is the reference group.

Daily Intention Enactment

The outcome variable measuring participants' daily enactment of the desired behavior (ie, water drinking) was subjected to a Poisson regression model with the same set of predictors as for *engagement*. Critically, we found that participants in the optimized gamification condition enacted the daily intention to

drink water more often (mean 14.71, SD 6.57 times) than either the participants in the baseline condition (mean 11.64, SD 5.43 times) or the participants in the control condition with reminders and feedback (mean 11.64, SD 6.38 times; [Table 4](#)). Furthermore, reminders and feedback without points did not result in more water drinking than the *baseline condition* (incidence rate ratio=1.03, 95% CI 0.90-1.17; $P=.70$).

Table 4. Predicting the number of days on which participants drank water (daily intention enactment) from their condition and preintervention measurements of habit-related control variables (n=118)^a.

Predictor	Incidence rate ratio (95% CI)	P value
Intercept	16.02 (13.93-18.41)	<.001
Baseline vs optimized gamification	0.78 (0.69-0.89)	.001
Reminders and feedback vs optimized gamification	0.80 (0.71-0.91)	<.001
Preintervention retrospective intention enactment	1.01 (0.97-1.05)	.64
History of repetition	0.91 (0.84-0.99)	.02
Automaticity	1.07 (0.99-1.15)	.10

^aNagelkerke $R^2=0.21$. The treatment group with optimal points is the reference group.

Self-Reported Habit Strength

To test whether optimized gamification promoted habit formation, we compared the change in the SRHI self-report measures of *automaticity* and *history of repetition* from pretest to posttest to follow-up time points among the 3 experimental conditions (Figure S1 in [Multimedia Appendix 3](#)) using a multilevel model with fixed-effects predictors for time point, 2 dummy codes for the experimental condition with the *optimized gamification condition* as the reference, and all pairwise *time*×*condition* interactions (Tables S1 and S2 in [Multimedia Appendix 3](#)). We found that, compared with the pretest time point, both measures of habit strength were significantly higher immediately after the intervention ($P<.001$ in all cases) and at follow-up ($P<.001$ in all cases). However, these effects were no larger in the *optimized gamification condition* than in the baseline condition ($P\geq.20$ in all cases) or the control condition with reminders and feedback ($P\geq.12$ in all cases).

Behavior After the Intervention

As a further test of whether the behavior change we observed during the intervention was maintained, we analyzed the number of times participants reported having enacted their intention in the week before the intervention versus the first week after the

intervention and the second week after the intervention (*retrospective intention enactment*; Figure S2 in [Multimedia Appendix 3](#)) using a multilevel model with fixed-effects predictors for time point, 2 dummy codes for the experimental condition with the *optimized gamification condition* as the reference, and all pairwise *time*×*condition* interactions (Table S3 in [Multimedia Appendix 3](#)). We found that, compared with the week before the intervention (mean 1.8, SD 2.1 times), participants enacted the desired behavior significantly more often after the intervention (1-week follow-up: mean 5.2, SD 1.9 times, $t_{299}=15.21$, and $P<.001$; 2-week follow-up: mean 5.1, SD 2.1 times, $t_{299}=14.86$, and $P<.001$). However, these effects were no larger in the optimized gamification condition than they were in the baseline condition ($P\geq.18$ in all cases) or the control condition with reminders and feedback ($P\geq.32$ in all cases).

Discussion

Principal Findings

In this study, we derived a mathematical principle for designing the point systems of gamified behavior change interventions. Our proof-of-concept study suggests that this principled

approach to gamifying behavior change can be beneficial. We found that our gamified behavior change chatbot fostered positive behavior change during the intervention. This is consistent with previous findings that goal setting, reinforcement, reminders, and self-monitoring are effective techniques for promoting behavior change [7,33,39,41,44].

Moreover, we found that the behavior change that occurred during the intervention was maintained in all 3 conditions. The elements that the behavior change interventions in all 3 groups shared were goal setting and self-monitoring. Therefore, goal setting and self-monitoring may be sufficient for sustained behavior change. Adding reinforcement to goal setting and self-monitoring was beneficial during the intervention, but the additional benefits of reinforcement ceased to be statistically significant ($P \geq .17$ in all cases) in the week following the intervention. However, as our study had a small sample size, this apparent discrepancy could be an artifact of us having used different methods to measure behavior change during versus after the intervention. During the intervention, we measured behavior change through daily self-reports. After the intervention, we asked participants to retrospectively report on their behavior in the previous week, which is less accurate because of participants' fallible memory, and complete self-report questionnaires about their perceived habit strength, which are less objective than reports on behavior. Consistent with the interpretation that our study had insufficient statistical power for detecting retention effects, the measures we used to assess the maintenance of behavior change consistently showed a nonsignificant trend in favor of optimized gamification (Figures S1 and S2 in [Multimedia Appendix 3](#)). Moreover, previous studies have found that gamification-induced behavior change can persist over extended periods [63].

Limitations

From a theoretical perspective, the main limitation of our study is that it did not compare the effectiveness of the points computed by our optimized gamification method with alternative point schemes. Previous work has found optimized gamification to be more effective than several heuristic methods for designing point systems in contexts in which people choose among several activities [51]. However, the decisions that people face in behavior change applications appear to be simpler. Therefore, it remains unclear how much of the beneficial effects of optimized gamification on behavior change are specific to optimized gamification. Relatedly, it remains unclear which property of the points generated through optimized gamification is responsible for their effects on behavior change. Future work could address these questions by comparing optimized gamification of behavior change with simpler, alternative feedback mechanisms such as always awarding the same number of points or a streak-based point system.

However, we did evaluate optimized gamification against rewarding each enactment of the desired behavior using the same positive feedback message and punishing each failure to enact the desired behavior using the same negative feedback message (control condition with reminders and feedback). From a reinforcement learning perspective, this condition is equivalent to always awarding the same number of points when the

behavior is enacted and always deducting the same number of points when the user fails to enact the behavior. We found that optimized gamification is more effective than this alternative feedback mechanism. This suggests that optimized gamification might be more effective than awarding the same number of points for each instance of the desired behavior. However, whether this interpretation is correct remains to be tested.

We illustrated the application of the general framework of optimized gamification to behavior change using a simple model of habit formation, which assumes that the user will indefinitely maintain the good habit once it has been established. This assumption is highly simplistic. In reality, maintenance is far from automatic. On the contrary, people may experience backsliding, and the strength of the habit may continue to wax and wane [64]. This could be captured by letting the process of deciding whether to perform the behavior continue indefinitely until the user dies. In this way, lapses could occur at any time and weaken the habit no matter how strong it is. In such a model, the health benefits of the behavior could be modeled explicitly in terms of its effects on a state variable that models the user's health status. Refining our method's model in this way would reduce the reinforcement for engaging in the behavior when the habit is weak and increase it when the habit is strong. This may make our method even more effective.

From an empirical perspective, the main limitation of the work presented in this paper is the relatively small sample size of our field experiment. Given that we collected <30 complete data sets per condition, the power of some of the statistical tests is relatively low. Therefore, the absence of significant differences in self-reported habit strength and retroactive intention enactment after the intervention does not provide strong evidence against the maintenance of the benefits of optimized gamification. Moreover, given that habit formation can take a very long time, our intervention may have been too short to fully capture the effects of the 3 different interventions on habit formation.

One flaw in our experiment was that some of the chatbot's messages were not worded in perfect, idiomatic English ([Multimedia Appendix 2](#)). We think that it is unlikely that our participants misunderstood any of the messages. However, it is possible that participants would have taken the messages more seriously if all of them had been written in perfect, idiomatic English. Another minor limitation of our chatbot is that its users started with a score of 0. Thus, if they failed to enact the desired behavior, their score fell to negative values, which might be demotivating. Therefore, future versions of our chatbot will award users a number of points (eg, 20) for setting the intention to build a good habit.

Finally, another weakness of our study design is that our intervention sought to strengthen water drinking in healthy people. Therefore, we cannot draw conclusions about the potential utility of our chatbot for clinical populations for which developing a healthy water-drinking habit might be crucial [3,4]. Moreover, it remains unclear whether our findings can be generalized to other habits that are more vital to people's health. In addition, our study was not specifically about water drinking as a weight loss strategy as only some participants tied water

drinking to their meals. However, we introduced a general method that can be used to improve digital interventions for many critical behavior change applications.

Comparison With Prior Work

This study builds on previous work on optimized gamification [51]. Optimized gamification has been previously applied to help people decide what to work on [51] and which goals to set [52,55,56,65]. Moreover, optimized gamification has also been applied to help people stay focused on their work [58]. However, the work presented in this paper is the first application of optimized gamification to support habit formation. Moreover, it is one of the first real-world applications of optimized gamification as most previous work was confined to controlled laboratory experiments.

As our chatbot combined 4 established behavior change techniques (ie, goal setting, reminders, support for self-monitoring, and reinforcement), its design and effectiveness are therefore consistent with several extant theories of behavior change [66]. In particular, our optimized gamification method is consistent with behavior change methods that acknowledge the importance of providing positive feedback on improvements in behavior [28,66-69].

The most similar gamified digital intervention for behavior change that we are aware of is the SMS text messaging-based WalkIT intervention for promoting physical exercise [7,33,70]. Participants of the WalkIT trial received SMS text messages with daily physical exercise goals. Physical exercise (walking) was measured using the accelerometers of their smartphones and reported to a server. Participants received feedback on whether they met the exercise goal via an SMS text message that included points that were converted into money. Depending on the stage of the experiment, the number of points for achieving a goal was either constant or determined at random. The number of points awarded for failing to achieve a goal was 0. In contrast, our optimized gamification method provides a principled way to choose the exact number of points that a person should be awarded for meeting their daily goal or lose for failing to meet that goal. Therefore, our method could be used to enhance the WalkIT intervention with a more principled way of choosing the number of points depending on the user's history of successful and unsuccessful goal achievement. Conversely, WalkIT has many sophisticated features that go beyond the chatbot we introduced here. This includes an algorithm for adaptive goal setting, automatically delivered financial incentives, and an evidence-based sequence of different reward schedules that differ in the probability that goal achievement will be rewarded and whether the magnitude of the reward will be fixed or random. Consistent with our finding that the more performance-contingent feedback of the optimized gamification condition was more effective than less informative feedback or no feedback, the WalkIT studies found that immediate, behavior-contingent reinforcement was more effective in promoting behavior change than delayed, behavior-independent reinforcement.

Another gamified digital intervention for supporting behavior change is Habitica. Habitica embeds working through one's to-do list into a role-playing game in which the user's character

can earn points by completing their daily to-dos. The points serve as an in-game currency that the player can use to buy weapons and armor for their avatar. Conversely, when the user does not complete a daily to-do, they lose points. As far as we can tell, the developers of Habitica chose the number of points the user gains for completing a to-do and the number of points they lose for failing to complete a daily to-do somewhat arbitrarily based on their intuitions. A recent study found that only 49% of Habitica's users rate its rewards as (rather) appropriate and that most experience counterproductive effects of Habitica's approach to gamification [17]. Given that optimized gamification was effective in our study and in previous studies, it is possible that redesigning Habitica's point system according to optimized gamification could alleviate some of the counterproductive effects of their users' experience.

The method introduced in this study mitigates the adverse effects of temporal discounting on people's health behavior [26]. Its approach is to add immediate rewards that are aligned with the behavior's long-term consequences for the user's health; that is, optimized gamification redesigns the decision environment so that people's shortsighted biases lead to decisions that are good for them in the long run [51]. Recent work on this topic introduced a computational model of intertemporal choice and applied it to compute personalized incentives for helping people more patiently work toward obtaining a larger reward later instead of abandoning the project in favor of a smaller immediate reward [71,72]. Similar to optimized gamification, their approach uses an MDP framework. However, their problem formulation and solution are different. The main difference lies in the application area. Sukumar et al [71] focused on canonical delay-of-gratification tasks, whereas we modeled habit formation. They tested their approach in online experiments in which participants played a *queue waiting game* and found that personalized incentives can increase people's patience while waiting in a simulated queue. In contrast, we conducted a field experiment on behavior change in which the incentives motivated people to act more farsightedly in the real world. Despite this critical difference, investigating whether modeling and measuring individual differences can be used to make optimized gamification more effective is an exciting direction for future work. Moreover, a computational model such as the one proposed by Sukumar et al [71] could be used to simulate the effects of alternative incentive schemes.

Previous work has found that drinking water before meals is an effective weight loss strategy for adults with obesity [3,4]. In the randomized controlled trial by Parretti et al [4], participants were instructed to use the water-drinking strategy in a face-to-face weight loss consultation. They did not receive any additional support in implementing this strategy. The chatbot we developed could be used to augment those weight loss consultations with a digital tool that helps people follow through on their resolutions. Alternatively, an appropriately adapted version of our chatbot could be used as a highly scalable, low-cost alternative to face-to-face weight loss consultations.

Conclusions

In conclusion, optimized gamification is a practically helpful theoretical principle for designing the point systems of (digital)

behavior change tools and interventions. It can be implemented in just a few lines of code, and the point values can be computed instantaneously. It can be applied to improve or augment many existing (digital) behavior change interventions and can also be used to create new ones. Thus, optimized gamification can help tackle many challenging behavior change problems using scalable digital interventions. Testing whether, when, and how

optimized gamification can make a positive difference in critical practical applications is an exciting direction for future research. A crucial next step will be to test our point system against simpler heuristic point systems for supporting behavior change. Moreover, our chatbot can be extended to support various health behaviors and other forms of positive behavior change.

Acknowledgments

This work was supported by grant 1757269 from the National Science Foundation.

Data Availability

The data generated in this study, the R scripts used to analyze the data, and the MATLAB (MathWorks) implementation of the optimized gamification method are available on the project's Open Science Framework repository [73].

Authors' Contributions

FL conceived and designed the study, developed the optimized gamification method, and wrote the manuscript. PZC designed and implemented the Good Habit Bot and co-designed and conducted the field experiment. MP conducted the data analysis. VA supported PZC in designing and conducting the field experiment and in the design of the Good Habit Bot. MT contributed to the implementation of the Good Habit Bot.

Conflicts of Interest

None declared.

Multimedia Appendix 1

An optimal feedback method for accelerating positive behavior change.

[\[DOCX File , 16 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Details about the Good Habit Bot.

[\[DOCX File , 17 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Details about the statistical results and supplementary analyses.

[\[DOCX File , 122 KB-Multimedia Appendix 3\]](#)

References

1. Physical activity. World Health Organization. Oct 5, 2022. URL: <https://www.who.int/news-room/fact-sheets/detail/physical-activity> [accessed 2024-02-08]
2. Bull ER, McCleary N, Li X, Dombrowski SU, Dusseldorp E, Johnston M. Interventions to promote healthy eating, physical activity and smoking in low-income groups: a systematic review with meta-analysis of behavior change techniques and delivery/context. *Int J Behav Med*. Dec 12, 2018;25(6):605-616. [FREE Full text] [doi: [10.1007/s12529-018-9734-z](https://doi.org/10.1007/s12529-018-9734-z)] [Medline: [30003476](https://pubmed.ncbi.nlm.nih.gov/30003476/)]
3. Daniels MC, Popkin BM. Impact of water intake on energy intake and weight status: a systematic review. *Nutr Rev*. Sep 2010;68(9):505-521. [FREE Full text] [doi: [10.1111/j.1753-4887.2010.00311.x](https://doi.org/10.1111/j.1753-4887.2010.00311.x)] [Medline: [20796216](https://pubmed.ncbi.nlm.nih.gov/20796216/)]
4. Parretti HM, Aveyard P, Blannin A, Clifford SJ, Coleman SJ, Roalfe A, et al. Efficacy of water preloading before main meals as a strategy for weight loss in primary care patients with obesity: RCT. *Obesity (Silver Spring)*. Sep 03, 2015;23(9):1785-1791. [FREE Full text] [doi: [10.1002/oby.21167](https://doi.org/10.1002/oby.21167)] [Medline: [26237305](https://pubmed.ncbi.nlm.nih.gov/26237305/)]
5. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games*. Dec 12, 2013;1(1):e3. [FREE Full text] [doi: [10.2196/games.3139](https://doi.org/10.2196/games.3139)] [Medline: [25658754](https://pubmed.ncbi.nlm.nih.gov/25658754/)]
6. Berardi V, Hovell M, Hurley JC, Phillips CB, Bellettiere J, Todd M, et al. Correction to: variable magnitude and frequency financial reinforcement is effective at increasing adults' free-living physical activity. *Perspect Behav Sci*. Sep 12, 2020;43(3):539-541. [FREE Full text] [doi: [10.1007/s40614-020-00254-y](https://doi.org/10.1007/s40614-020-00254-y)] [Medline: [33031485](https://pubmed.ncbi.nlm.nih.gov/33031485/)]

7. Adams MA, Todd M, Angadi SS, Hurley JC, Stecher C, Berardi V, et al. Adaptive goals and reinforcement timing to increase physical activity in adults: a factorial randomized trial. *Am J Prev Med.* Feb 2022;62(2):e57-e68. [FREE Full text] [doi: [10.1016/j.amepre.2021.09.014](https://doi.org/10.1016/j.amepre.2021.09.014)] [Medline: [35000693](https://pubmed.ncbi.nlm.nih.gov/35000693/)]
8. Berardi V, Bellettiere J, Nguyen B, Klepeis NE, Hughes SC, Adams MA, et al. Assessing reinforcing versus aversive consequences in a real-time secondhand smoke intervention. *Transl Behav Med.* Aug 13, 2021;11(8):1558-1566. [FREE Full text] [doi: [10.1093/tbm/ibab004](https://doi.org/10.1093/tbm/ibab004)] [Medline: [33823045](https://pubmed.ncbi.nlm.nih.gov/33823045/)]
9. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. *J Biomed Inform.* Jul 2017;71:31-48. [FREE Full text] [doi: [10.1016/j.jbi.2017.05.011](https://doi.org/10.1016/j.jbi.2017.05.011)] [Medline: [28536062](https://pubmed.ncbi.nlm.nih.gov/28536062/)]
10. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. *Internet Interv.* Nov 2016;6:89-106. [FREE Full text] [doi: [10.1016/j.invent.2016.10.002](https://doi.org/10.1016/j.invent.2016.10.002)] [Medline: [30135818](https://pubmed.ncbi.nlm.nih.gov/30135818/)]
11. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. *Int J Med Inform.* Dec 2016;96:62-70. [doi: [10.1016/j.ijmedinf.2016.02.006](https://doi.org/10.1016/j.ijmedinf.2016.02.006)] [Medline: [26944611](https://pubmed.ncbi.nlm.nih.gov/26944611/)]
12. Schmidt-Kraepelin M, Warsinsky S, Thiebes S, Sunyaev A. The role of gamification in health behavior change: a review of theory-driven studies. In: *Proceedings of the 53rd Hawaii International Conference on System Sciences.* 2020. Presented at: HICSS 2020; January 7-10, 2020; Maui, Hawaii. URL: <https://scholarspace.manoa.hawaii.edu/items/3e3f2688-4f1c-42a2-840e-ccfedcd85e8> [doi: [10.24251/hicss.2020.155](https://doi.org/10.24251/hicss.2020.155)]
13. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments.* 2011. Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland. [doi: [10.1145/2181037.2181040](https://doi.org/10.1145/2181037.2181040)]
14. Werbach K, Hunter D. *For the Win: The Power of Gamification and Game Thinking in Business, Education, Government, and Social Impact.* Philadelphia, PA. Wharton School Press; 2020.
15. Zichermann G, Cunningham C. *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps.* Sebastopol, CA. O'Reilly Media; 2011.
16. Toda AM, Valle PH, Isotani S. The dark side of gamification: an overview of negative effects of gamification in education. In: *Proceedings of the First International Workshop on Social, Semantic, Adaptive and Gamification Techniques and Technologies for Distance Learning, HEFA 2017.* 2017. Presented at: First International Workshop on Social, Semantic, Adaptive and Gamification Techniques and Technologies for Distance Learning, HEFA 2017; March 20-24, 2017; Maceió, Brazil. URL: <https://link.springer.com/book/10.1007/978-3-319-97934-2> [doi: [10.1007/978-3-319-97934-2_9](https://doi.org/10.1007/978-3-319-97934-2_9)]
17. Diefenbach S, Müssig A. Counterproductive effects of gamification: an analysis on the example of the gamified task manager Habitica. *Int J Hum Comput Stud.* Jul 2019;127:190-210. [doi: [10.1016/j.ijhcs.2018.09.004](https://doi.org/10.1016/j.ijhcs.2018.09.004)]
18. Nicholson S. A RECIPE for meaningful gamification. In: Reiners T, Wood L, editors. *Gamification in Education and Business.* Cham, Switzerland. Springer; 2015.
19. Dolan RJ, Dayan P. Goals and habits in the brain. *Neuron.* Oct 16, 2013;80(2):312-325. [FREE Full text] [doi: [10.1016/j.neuron.2013.09.007](https://doi.org/10.1016/j.neuron.2013.09.007)] [Medline: [24139036](https://pubmed.ncbi.nlm.nih.gov/24139036/)]
20. Baumeister RF, Heatherton TF. Self-regulation failure: an overview. *Psychol Inq.* Jan 1996;7(1):1-15. [doi: [10.1207/s15327965pli0701_1](https://doi.org/10.1207/s15327965pli0701_1)]
21. Critchfield TS, Kollins SH. Temporal discounting: basic research and the analysis of socially important behavior. *J Appl Behav Anal.* Feb 27, 2001;34(1):101-122. [FREE Full text] [doi: [10.1901/jaba.2001.34-101](https://doi.org/10.1901/jaba.2001.34-101)] [Medline: [11317983](https://pubmed.ncbi.nlm.nih.gov/11317983/)]
22. O'Donoghue S, Rabin M. Doing it now or later. *Am Econ Rev.* 1999;89(1):103-124. [FREE Full text] [doi: [10.1257/aer.89.1.103](https://doi.org/10.1257/aer.89.1.103)]
23. Berns GS, Laibson D, Loewenstein G. Intertemporal choice--toward an integrative framework. *Trends Cogn Sci.* Nov 2007;11(11):482-488. [doi: [10.1016/j.tics.2007.08.011](https://doi.org/10.1016/j.tics.2007.08.011)] [Medline: [17980645](https://pubmed.ncbi.nlm.nih.gov/17980645/)]
24. Steel P, König CJ. Integrating theories of motivation. *Acad Manage Rev.* Oct 2006;31(4):889-913. [doi: [10.2307/20159257](https://doi.org/10.2307/20159257)]
25. Hennessy EA, Johnson BT, Acabchuk RL, McCloskey K, Stewart-James J. Self-regulation mechanisms in health behavior change: a systematic meta-review of meta-analyses, 2006-2017. *Health Psychol Rev.* Mar 12, 2020;14(1):6-42. [FREE Full text] [doi: [10.1080/17437199.2019.1679654](https://doi.org/10.1080/17437199.2019.1679654)] [Medline: [31662031](https://pubmed.ncbi.nlm.nih.gov/31662031/)]
26. Story GW, Vlaev I, Seymour B, Darzi A, Dolan RJ. Does temporal discounting explain unhealthy behavior? a systematic review and reinforcement learning perspective. *Front Behav Neurosci.* Mar 12, 2014;8:76. [FREE Full text] [doi: [10.3389/fnbeh.2014.00076](https://doi.org/10.3389/fnbeh.2014.00076)] [Medline: [24659960](https://pubmed.ncbi.nlm.nih.gov/24659960/)]
27. Rebar AL, Gardner B. Habit formation and behavior change. In: *Oxford Research Encyclopedia of Psychology.* Oxford, UK. Oxford University Press; 2019.
28. Skinner BF. *Science And Human Behavior.* New York, NY. Macmillan Publishers; 1953.
29. Staddon JE, Cerutti DT. Operant conditioning. *Annu Rev Psychol.* 2003;54:115-144. [FREE Full text] [doi: [10.1146/annurev.psych.54.101601.145124](https://doi.org/10.1146/annurev.psych.54.101601.145124)] [Medline: [12415075](https://pubmed.ncbi.nlm.nih.gov/12415075/)]
30. Miller KJ, Shenhav A, Ludvig EA. Habits without values. *Psychol Rev.* Mar 2019;126(2):292-311. [FREE Full text] [doi: [10.1037/rev0000120](https://doi.org/10.1037/rev0000120)] [Medline: [30676040](https://pubmed.ncbi.nlm.nih.gov/30676040/)]

31. Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: a meta-analysis. *Addiction*. Nov 10, 2006;101(11):1546-1560. [doi: [10.1111/j.1360-0443.2006.01581.x](https://doi.org/10.1111/j.1360-0443.2006.01581.x)] [Medline: [17034434](https://pubmed.ncbi.nlm.nih.gov/17034434/)]
32. Higgins ST, Budney AJ, Bickel WK, Foerg FE, Donham R, Badger GJ. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. *Arch Gen Psychiatry*. Jul 01, 1994;51(7):568-576. [doi: [10.1001/archpsyc.1994.03950070060011](https://doi.org/10.1001/archpsyc.1994.03950070060011)] [Medline: [8031230](https://pubmed.ncbi.nlm.nih.gov/8031230/)]
33. Berardi V, Hovell M, Hurley JC, Phillips CB, Bellettiere J, Todd M, et al. Variable magnitude and frequency financial reinforcement is effective at increasing adults' free-living physical activity. *Perspect Behav Sci*. Sep 05, 2020;43(3):515-538. [FREE Full text] [doi: [10.1007/s40614-019-00241-y](https://doi.org/10.1007/s40614-019-00241-y)] [Medline: [33029578](https://pubmed.ncbi.nlm.nih.gov/33029578/)]
34. Pfund RA, Ginley MK, Rash CJ, Zajac K. Contingency management for treatment attendance: a meta-analysis. *J Subst Abuse Treat*. Feb 2022;133:108556. [FREE Full text] [doi: [10.1016/j.jsat.2021.108556](https://doi.org/10.1016/j.jsat.2021.108556)] [Medline: [34210566](https://pubmed.ncbi.nlm.nih.gov/34210566/)]
35. Griffith JD, Rowan-Szal GA, Roark RR, Simpson DD. Contingency management in outpatient methadone treatment: a meta-analysis. *Drug Alcohol Depend*. Feb 01, 2000;58(1-2):55-66. [doi: [10.1016/s0376-8716\(99\)00068-x](https://doi.org/10.1016/s0376-8716(99)00068-x)] [Medline: [10669055](https://pubmed.ncbi.nlm.nih.gov/10669055/)]
36. Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction*. Feb 2006;101(2):192-203. [doi: [10.1111/j.1360-0443.2006.01311.x](https://doi.org/10.1111/j.1360-0443.2006.01311.x)] [Medline: [16445548](https://pubmed.ncbi.nlm.nih.gov/16445548/)]
37. Roll JM, Higgins ST, Badger GJ. An experimental comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. *J Appl Behav Anal*. Feb 27, 1996;29(4):495-504; quiz 504. [FREE Full text] [doi: [10.1901/jaba.1996.29-495](https://doi.org/10.1901/jaba.1996.29-495)] [Medline: [8995832](https://pubmed.ncbi.nlm.nih.gov/8995832/)]
38. Taj F, Klein MC, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. *JMIR Mhealth Uhealth*. Dec 13, 2019;7(12):e13311. [FREE Full text] [doi: [10.2196/13311](https://doi.org/10.2196/13311)] [Medline: [31833836](https://pubmed.ncbi.nlm.nih.gov/31833836/)]
39. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. *J Adolesc Health*. Dec 2017;61(6):669-677. [FREE Full text] [doi: [10.1016/j.jadohealth.2017.05.024](https://doi.org/10.1016/j.jadohealth.2017.05.024)] [Medline: [28822682](https://pubmed.ncbi.nlm.nih.gov/28822682/)]
40. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. *Am J Public Health*. Apr 2011;101(4):751-758. [FREE Full text] [doi: [10.2105/AJPH.2010.194381](https://doi.org/10.2105/AJPH.2010.194381)] [Medline: [21330590](https://pubmed.ncbi.nlm.nih.gov/21330590/)]
41. Eckerstorfer LV, Tanzer NK, Vogrincic-Haselbacher C, Kedia G, Brohmer H, Dinslaken I, et al. Key elements of mHealth interventions to successfully increase physical activity: meta-regression. *JMIR Mhealth Uhealth*. Nov 12, 2018;6(11):e10076. [FREE Full text] [doi: [10.2196/10076](https://doi.org/10.2196/10076)] [Medline: [30425028](https://pubmed.ncbi.nlm.nih.gov/30425028/)]
42. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: a meta - analysis of effects and processes. *Adv Exp Soc Psychol*. 2006;38:119. [doi: [10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)]
43. Gollwitzer PM. Implementation intentions: strong effects of simple plans. *Am Psychol*. Jul 1999;54(7):493-503. [doi: [10.1037/0003-066X.54.7.493](https://doi.org/10.1037/0003-066X.54.7.493)]
44. Figueroa CA, Deliu N, Chakraborty B, Modiri A, Xu J, Aggarwal J, et al. Daily motivational text messages to promote physical activity in university students: results from a microrandomized trial. *Ann Behav Med*. Feb 11, 2022;56(2):212-218. [doi: [10.1093/abm/kaab028](https://doi.org/10.1093/abm/kaab028)] [Medline: [33871015](https://pubmed.ncbi.nlm.nih.gov/33871015/)]
45. Hamari J, Koivisto J, Sarsa H. Does gamification work? -- a literature review of empirical studies on gamification. In: *Proceedings of the 47th Hawaii International Conference on System Sciences*. 2014. Presented at: HICSS 2014; January 6-9, 2014; Waikoloa, HI. [doi: [10.1109/hicss.2014.377](https://doi.org/10.1109/hicss.2014.377)]
46. Mekler ED, Brühlmann F, Opwis K, Tuch AN. Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of common gamification elements. In: *Proceedings of the First International Conference on Gameful Design, Research, and Applications*. 2013. Presented at: Gamification '13; October 2-4, 2013; Toronto, Ontario. [doi: [10.1145/2583008.2583017](https://doi.org/10.1145/2583008.2583017)]
47. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Comput Hum Behav*. Jun 2017;71:525-534. [doi: [10.1016/j.chb.2015.08.048](https://doi.org/10.1016/j.chb.2015.08.048)]
48. Antonaci A, Klemke R, Specht M. The effects of gamification in online learning environments: a systematic literature review. *Informatics*. Aug 12, 2019;6(3):32. [doi: [10.3390/informatics6030032](https://doi.org/10.3390/informatics6030032)]
49. Schmidt-Kraepelin M, Thiebes S, Stepanovic S, Mettler T, Sunyaev A. Gamification in health behavior change support systems - a synthesis of unintended side effects. In: *Proceedings of the 14th International Conference on Wirtschaftsinformatik*. 2019. Presented at: WI2019; February 24-27, 2019; Siegen, Germany.
50. Nacke LE, Deterding S. The maturing of gamification research. *Comput Hum Behav*. Jun 2017;71:450-454. [doi: [10.1016/j.chb.2016.11.062](https://doi.org/10.1016/j.chb.2016.11.062)]
51. Lieder F, Chen OX, Krueger PM, Griffiths TL. Cognitive prostheses for goal achievement. *Nat Hum Behav*. Oct 19, 2019;3(10):1096-1106. [doi: [10.1038/s41562-019-0672-9](https://doi.org/10.1038/s41562-019-0672-9)] [Medline: [31427788](https://pubmed.ncbi.nlm.nih.gov/31427788/)]
52. Consul S, Stojcheski J, Lieder F. Leveraging AI for effective to-do list gamification. In: *Proceedings of the 5th International Workshop "Gam-R – Gamification Reloaded"*. 2022. Presented at: 5th International Workshop "Gam-R – Gamification Reloaded"; September 04, 2022; Darmstadt, Germany. URL: <https://dl.gi.de/items/b17fae53-fdd1-4d91-a122-7df976acc37d> [doi: [10.1515/icom-2021-0025](https://doi.org/10.1515/icom-2021-0025)]

53. Steel P. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychol Bull.* Jan 2007;133(1):65-94. [doi: [10.1037/0033-2909.133.1.65](https://doi.org/10.1037/0033-2909.133.1.65)] [Medline: [17201571](https://pubmed.ncbi.nlm.nih.gov/17201571/)]
54. Murayama K. A reward-learning framework of knowledge acquisition: an integrated account of curiosity, interest, and intrinsic-extrinsic rewards. *Psychol Rev.* Jan 2022;129(1):175-198. [doi: [10.1037/rev0000349](https://doi.org/10.1037/rev0000349)] [Medline: [35099213](https://pubmed.ncbi.nlm.nih.gov/35099213/)]
55. Xu L, Wirzberger M, Lieder F. How should we incentivize learning? an optimal feedback mechanism for educational games and online courses. In: *Proceedings of the 41st Annual Meeting of the Cognitive Science Society.* 2019. Presented at: CogSci 2019; July 24-27, 2019; Montreal, Canada.
56. Pauly R, Heindrich L, Amo V, Lieder F. What to learn next? aligning gamification rewards to long-term goals using reinforcement learning. In: *Proceedings of the Multi-disciplinary Conference on Reinforcement Learning and Decision Making 2022.* 2022. Presented at: RLDLM; June 8-11, 2022; Providence, RI. [doi: [10.1002/9781118266502.ch3](https://doi.org/10.1002/9781118266502.ch3)]
57. Callaway F, Jain YR, van Opheusden B, Das P, Iwama G, Gul S, et al. Leveraging artificial intelligence to improve people's planning strategies. *Proc Natl Acad Sci U S A.* Mar 22, 2022;119(12):e2117432119. [FREE Full text] [doi: [10.1073/pnas.2117432119](https://doi.org/10.1073/pnas.2117432119)] [Medline: [35294284](https://pubmed.ncbi.nlm.nih.gov/35294284/)]
58. Wirzberger M, Lado A, Eckerstorfer L, Oreshnikov I, Passy J, Stock A, et al. How to navigate everyday distractions: leveraging optimal feedback to train attention control. In: *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society.* 2020. Presented at: CogSci 2020; July 29-August 1, 2020; Online. [doi: [10.21203/rs.3.rs-3355487/v1](https://doi.org/10.21203/rs.3.rs-3355487/v1)]
59. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction.* Cambridge, MA. MIT Press; 2018.
60. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game on? Smoking cessation through the gamification of mHealth: a longitudinal qualitative study. *JMIR Serious Games.* Oct 24, 2016;4(2):e18. [FREE Full text] [doi: [10.2196/games.5678](https://doi.org/10.2196/games.5678)] [Medline: [27777216](https://pubmed.ncbi.nlm.nih.gov/27777216/)]
61. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. *J Appl Soc Psychol.* Jul 31, 2006;33(6):1313-1330. [doi: [10.1111/j.1559-1816.2003.tb01951.x](https://doi.org/10.1111/j.1559-1816.2003.tb01951.x)]
62. Prentice M, Lieder F. Habit formation REG. *Open Science Framework.* Jun 8, 2022. URL: <https://osf.io/h7jn8> [accessed 2024-02-09]
63. Harris MA. Maintenance of behaviour change following a community-wide gamification based physical activity intervention. *Prev Med Rep.* Mar 2019;13:37-40. [FREE Full text] [doi: [10.1016/j.pmedr.2018.11.009](https://doi.org/10.1016/j.pmedr.2018.11.009)] [Medline: [30510892](https://pubmed.ncbi.nlm.nih.gov/30510892/)]
64. Dunton GF, Leventhal AM, Rebar AL, Gardner B, Intille SS, Rothman AJ. Towards consensus in conceptualizing and operationalizing physical activity maintenance. *Psychol Sport Exerc.* Jul 2022;61:102214. [doi: [10.1016/j.psychsport.2022.102214](https://doi.org/10.1016/j.psychsport.2022.102214)]
65. Lieder F, Chen PZ, Stojcheski J, Consul S, Pammer-Schindler V. A cautionary tale about AI-generated goal suggestions. In: *Proceedings of the Mensch und Computer 2022.* 2022. Presented at: MuC '22; September 4-7, 2022; Darmstadt, Germany. URL: <https://dl.acm.org/doi/fullHtml/10.1145/3543758.3547539> [doi: [10.1145/3543758.3547539](https://doi.org/10.1145/3543758.3547539)]
66. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath V, editors. *Health Behavior: Theory, Research, and Practice.* San Francisco, CA. Jossey-Bass Inc; 2015.
67. Hovell M, Wahlgren D, Adams M. The logical and empirical basis for the behavioral ecological model. In: DiClemente RJ, Crosby RA, Kegler MC, editors. *Emerging Theories in Health Promotion Practice and Research.* San Francisco, CA. Jossey-Bass Inc; 2009.
68. Sallis JF, Owen N, Fisher EB. Ecological models of health behavior. In: Glanz K, Rimer BK, Viswanath K, editors. *Health Behavior and Health Education: Theory, Research, and Practice.* San Francisco, CA. Jossey-Bass Inc; 2008.
69. Bandura A. *Self-Efficacy: The Exercise of Control.* New York, NY. Worth Publishers; 1997.
70. Adams MA, Hurley JC, Phillips CB, Todd M, Angadi SS, Berardi V, et al. Rationale, design, and baseline characteristics of WalkIT Arizona: a factorial randomized trial testing adaptive goals and financial reinforcement to increase walking across higher and lower walkable neighborhoods. *Contemp Clin Trials.* Jun 2019;81:87-101. [FREE Full text] [doi: [10.1016/j.cct.2019.05.001](https://doi.org/10.1016/j.cct.2019.05.001)] [Medline: [31063868](https://pubmed.ncbi.nlm.nih.gov/31063868/)]
71. Sukumar S, Ward AF, Elliott-Williams C, Hakimi S, Mozer MC. Overcoming temptation: incentive design for intertemporal choice. *arXiv. Preprint posted online March 11, 2022.* 2023. [FREE Full text]
72. Mozer MC, Sukumar S, Elliott-Williams C, Hakimi S, Ward AF. Overcoming temptation: incentive design for intertemporal choice. In: *Proceedings of the 3rd Multidisciplinary Conference on Reinforcement Learning and Decision Making.* 2017. Presented at: RLDLM '17; June 11-14, 2017; Ann Arbor, MI.
73. Prentice M, Lieder F, Chen PZ. Habit formation REG. *Open Science Framework.* Jun 5, 2022. URL: <https://osf.io/5qg2n> [accessed 2024-02-08]

Abbreviations

MDP: Markov decision process

SRHI: Self-Report Habit Index

Edited by G Eysenbach, T Leung; submitted 30.09.22; peer-reviewed by V Berardi, M Kapsetaki; comments to author 25.01.23; revised version received 12.06.23; accepted 31.08.23; published 22.03.24

Please cite as:

Lieder F, Chen PZ, Prentice M, Amo V, Tošić M

Gamification of Behavior Change: Mathematical Principle and Proof-of-Concept Study

JMIR Serious Games 2024;12:e43078

URL: <https://games.jmir.org/2024/1/e43078>

doi: [10.2196/43078](https://doi.org/10.2196/43078)

PMID: [38517466](https://pubmed.ncbi.nlm.nih.gov/38517466/)

©Falk Lieder, Pin-Zhen Chen, Mike Prentice, Victoria Amo, Mateo Tošić. Originally published in JMIR Serious Games (<https://games.jmir.org>), 22.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on <https://games.jmir.org>, as well as this copyright and license information must be included.