Like Adding a Small Weight to a Scale About to Tip: Personalizing Micro-Financial Incentives for Digital Wellbeing

Sueun Jang
School of Computing
KAIST
Daejeon, Republic of Korea
sueun.jang@kaist.ac.kr

Youngseok Seo
School of Computing
KAIST
Daejeon, Republic of Korea
ysseo0910@kaist.ac.kr

Woohyeok Choi*
Dept. of Computer Science and Engineering
Kangwon National University
Chuncheon-si, Republic of Korea
woohyeok.choi@kangwon.ac.kr

Uichin Lee*
School of Computing
KAIST
Daejeon, Republic of Korea
uclee@kaist.edu

Abstract

Personalized behavior change interventions can be effective as they dynamically adapt to an individual's context. Financial incentives, a commonly used intervention in commercial applications and policy-making, offer a mechanism for creating personalized micro-interventions that are both quantifiable and amenable to systematic evaluation. However, the effectiveness of such personalized micro-financial incentives in real-world settings remains largely unexplored. In this study, we propose a personalization strategy that dynamically adjusts the amount of micro-financial incentives to promote smartphone use regulation and explore its efficacy and user experience through a four-week, in-the-wild user study. The results demonstrate that the proposed method is highly cost-effective without compromising intervention effectiveness. Based on these findings, we discuss the role of micro-financial incentives in enhancing awareness, design considerations for personalized micro-financial incentive systems, and their potential benefits and limitations concerning motivation change.

CCS Concepts

• **Human-centered computing → Empirical studies in ubiquitous and mobile computing**.

Keywords

digital wellbeing, micro-intervention, micro-financial incentive, behavior change, personalization

ACM Reference Format:
Sueun Jang, Youngseok Seo, Woohyeok Choi, and Uichin Lee. 2025. Like Adding a Small Weight to a Scale About to Tip: Personalizing Micro-Financial Incentives for Digital Wellbeing. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3706598.3714208

*Corresponding authors

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3714208

1 Introduction

The proliferation of smart devices and sensors has enabled fine-grained tracking of health and wellbeing behaviors while simultaneously increasing accessibility to care. This trend has spurred opportunities to design micro-interventions, which are shorter and highly focused interventions that offer small units of digital treatments to achieve specific wellness objectives in daily life [35, 75]. Micro-interventions, often termed just-in-time adaptive interventions (JITAIs) [83] or ecological momentary interventions [43], have been employed across various health and wellbeing domains, including mitigating prolonged sedentariness [15], managing stress [45, 87], fostering positive mood [75], and promoting digital wellbeing [80].

Prior research has identified two principal design considerations for enhancing user engagement with micro-interventions: delivery timing and intervention content [45]. Optimizing delivery timing involves ensuring that intervention content is delivered when users are most capable of processing and engaging with it. Prior studies have explored various concepts to achieve this, including interruptibility [129], availability [105], reachability [118], and receptivity [16]. Tailoring intervention content involves providing personalized content adapted to each individual's preferences, contexts, performance, or characteristics. For example, micro-interventions can deliver tailored content that considers an individual's current context [56]. Furthermore, the optimal content that maximizes beneficial effects on health outcomes can be automatically selected and recommended [4, 90, 97], or users can be empowered to choose their preferred health content [22].

While the appropriate design of delivery timing and content for micro-interventions has shown promise for improving user engagement, we posit that feedback or incentives aimed at reinforcing users' successful behavior change also offer significant potential for personalization. A common form of non-financial feedback is the badge, widely used to boost user participation and achievement in online communities [5, 128], education [6, 40], and gaming [36]. Studies have demonstrated that the effectiveness of badges in influencing user behavior varies based on design choices [49] and user demographics [128]. Similarly, in contingency management–a behavioral intervention approach that elicits positive behavior changes by providing external rewards (e.g., financial incentives) [93]–individuals' characteristics or socioeconomic status can influence the effectiveness of incentives [71, 78], suggesting that even financial incentives may benefit from personalization.

This study focuses on financial incentives as a mechanism for personalized micro-interventions, enabling quantitative comparison and systematic evaluation across different micro-intervention variations. However, implementing financial incentives as interventions requires careful consideration for several reasons. First, the financial resources provided by program-hosting institutions are inherently finite, necessitating budget constraints. Second, financial incentives may have the unintended side effect of undermining intrinsic motivation, known as the overjustification or crowding-out effect [23, 33, 34]. Third, the inherent complexity of financial incentives, involving various factors, makes it challenging to predict their precise impact on behavior change success [2]. Given these considerations, it is crucial to design financial incentives that can sufficiently engage individuals in the behavior change process while simultaneously minimizing the total financial outlay and gradually phasing out the external motivator.

To this end, we demonstrate a working example of a personalized behavior change intervention using financial incentives that dynamically adjust incentive amounts based on user behavior, specifically targeting digital wellbeing. Digital wellbeing has gained significant attention due to growing concerns about digital overuse, and consequently, various behavioral interventions have been explored to regulate digital overuse [88, 99]. Moreover, digital behavior is convenient to track and monitor, making it an ideal target domain for our study, which requires fine-grained tracking and timely intervention delivery.

Specifically, we propose a personalized micro-financial incentive strategy that balances behavior change promotion with overall incentive costs. Our personalization algorithm frames the problem of determining the optimal amount of micro-financial incentives as a multi-armed bandit problem, where different incentive amounts (i.e., arms) are associated with varying probabilities of behavior change success. Furthermore, our algorithm selects the optimal incentive amount designed to optimize dual objectives–maximizing the likelihood of behavior change and minimizing the expected cost–utilizing the Pareto front for efficient selection. Based on these problem definitions, our study deploys a mobile application in-the-wild with 72 participants over four weeks to investigate the efficacy and user experience of the proposed incentive strategy. Through quantitative and qualitative analysis of the user study, we demonstrate the cost-effectiveness of the proposed method while maintaining comparable performance in promoting smartphone usage regulation. Finally, we discuss the role of micro-financial incentives in enhancing awareness, design considerations for personalized micro-financial incentive systems, and their potential benefits and limitations concerning motivation change.

The key contributions of our study are as follows:

- We designed and implemented WellbeingWallet, a novel micro-intervention system for digital wellbeing that provided personalized financial incentives by leveraging both cognitive approaches (i.e., raising awareness through time-boxing and notifications) and behavioral approaches (i.e., providing automated missions and financial incentives) to promote self-regulated smartphone use.
- We conducted a four-week, in-the-wild study with 72 participants to evaluate the effectiveness and user experience of our personalized micro-financial incentive program, comparing it to non-personalized incentive programs (i.e., programs employing random and fixed incentive amounts).
- We provided design implications for designing computerized behavior change programs based on our empirical findings, particularly those incorporating personalized micro-financial incentives.

2 Related Work

2.1 Adaptive Intervention Techniques for Behavior Change in HCI

With the increasing number of everyday devices such as smartphones and wearables, behavior change researchers have moved beyond traditional methods, such as sending written letters [18] or SMS messages [30, 110], and started incorporating adaptivity and responsiveness into behavioral interventions. These responsive, interactive systems have enabled sophisticated adaptive interventions that provide users with just-in-time interventions tailored to their individual status and needs [43, 123]. To design such adaptive intervention systems, the HCI community has explored diverse techniques along four major dimensions: decision points (i.e., identifying the moment of decision-making), treatment options (i.e., providing a range of available treatments), tailoring variables (i.e., deciding what and how to measure from users to guide personalized interventions), and decision rules (i.e., determining specific treatment options based on the tailoring variables) [3]. For example, HeartSteps [63] proposed a personalization technique to identify opportune moments for delivering just-in-time physical activity suggestions by collecting mobile data such as location, temperature, step count, app engagement, and notification fatigue. Similarly, Time2Stop [86] introduced a machine learning model that determines opportune moments to provide tailored treatment options in a JITAI manner for digital wellbeing. PopTherapy [87] developed a personalization technique to decide treatment options based on tailoring variables by collecting user and contextual data. MyBehavior [97] investigated decision rules to control treatment options for personalization purposes by collecting activity tracking data and picture-based food-logging data. Smart-T alcohol [122] created personalized treatment messages by assessing personal risk for alcohol misuse from smartphone data.

Such JITAIs have shown promising results, particularly in health-related behavior change such as weight management [32] and smoking cessation [127], during both the intervention and post-intervention periods, compared to non-JITAI treatments [83]. Although studies have indicated that various extrinsic motivators, such as financial incentives [47], badge awards [52], and social interactions [59], can significantly impact behavior change and that these motivators need to be deliberately designed to fit individual needs and contexts, most research has adopted generalized approaches [29, 46, 59, 88, 120] rather than personalized, adaptive approaches. Only a few studies have attempted to leverage extrinsic motivators with high fidelity for adaptive, personalized interventions [15]. Therefore, our study aims to investigate this understudied approach that conditions treatment options at micro-levels to provide adaptive and personalized interventions.

2.2 Financial Incentives for Behavior Change

Financial incentives have emerged as a powerful tool for driving behavior change, particularly in the domains of public health and wellbeing [96]. They have been widely adopted in policy initiatives aimed at promoting healthier lifestyles and reducing societal burden associated with behaviors such as obesity [31], alcohol addiction [13], and smoking [91]. Through successful employments, financial incentives have demonstrated their effectiveness in healthcare settings, such as reinforcing positive habits [65], increasing medication adherence [25], and mitigating harmful behaviors such as substance misuse [112]. Furthermore, numerous studies have indicated that even modest financial incentives can positively influence short-term health-related behaviors [37, 84, 114, 121].

However, the design of effective financial incentive programs require careful consideration to avoid unintended consequences [14]. First, the relationship between the magnitude of the financial incentive and the resulting behavior change is complex and often unpredictable. Although traditional economic principles of marginal utility suggest diminishing returns with increasing incentive size, in practice, the decline in utility in behavior change contexts is often steeper and more variable than models predict [121]. Second, individual responses to financial incentives vary considerably based on personal characteristics, including gender [21], socioeconomic background, and sensitivity to monetary rewards [39]. This heterogeneity underscores the need for personalized approaches in designing financial interventions; a one-size-fits-all strategy is unlikely to be effective across diverse populations. Tailoring interventions to individual needs and preferences is crucial for maximizing their impact. Third, the financial sustainability of programs that rely on substantial rewards can be a significant concern [94]. Fourth, previous studies have suggested that extrinsic motivation driven by financial incentives can potentially crowd out intrinsic motivation [103] (i.e., a decline in the desired behavior once the incentive is removed). However, the crowding-out effect appears to be context-dependent and influenced by individual factors and situational characteristics [11]. Therefore, a personalized approach is crucial when employing financial incentives for behavior change.

Prior research has also investigated the effectiveness of financial incentives at different levels of granularity. These studies have found that applying financial incentives to fine-grained behaviors (e.g., micro-activity-based billing, adjusting incentives proportionally to the effort required for the behavior) is more effective than applying them to coarse-grained behaviors or solely to outcome-based measures [82, 126]. This fine-grained approach fosters a tighter link between positive behaviors and their associated rewards, promoting more sustainable behavior change [108]. Building on this concept, recent research has explored adaptive, fine-grained financial incentive strategies, such as BeActive [113] and StandUp [15], to encourage physical activity through the use of micro-financial incentives.

2.3 Digital Wellbeing Interventions

As concerns about the overuse of technology have escalated [80] such as increased stress, anxiety [119], and diminished interpersonal relationships [26, 130], there has been a growing interest in the development of sophisticated digital wellbeing techniques [76]. Furthermore, the convenience of tracking and monitoring digital behavior has facilitated experimentation with various intervention types. These interventions generally fall into one of the four categories [124]: self-monitoring (i.e., encouraging users to track their digital usage to increase awareness) [7, 100], reminders (i.e., notifying users to prompt them to take breaks or reduce smartphone use) [44, 88, 124], interaction friction (i.e., adding barriers or extra steps into digital interactions to curb excessive use) [85, 89, 125], and lockout (i.e., restricting access to certain apps or features to control smartphone use) [53, 54, 57, 58]. While these approaches have demonstrated promise, both self-monitoring and reminder-based interventions can be easily ignored, making it difficult for users to modify their problematic behaviors [80]. Conversely, while interaction friction and lockout interventions are more difficult to ignore [80], they can frustrate users when they try to use smartphones [54]. To address these limitations, incorporating financial incentives into self-monitoring and reminder mechanisms can offer a potential solution for enhancing the effectiveness of the interventions–e.g., by offering tangible rewards, reminders can become more engaging and less likely to be ignored or dismissed. By harnessing the motivational power of financial incentives to improve user responsiveness and adherence, this approach can address a critical weakness of traditional reminder systems. Thus, our study aims to explore a more effective, tailored intervention for digital wellbeing by combining micro-financial incentives with reminders.

3 Intervention Design

3.1 Design Rationale

3.1.1 Timebox-Based Behavioral Missions. Our intervention employed an hour-based micro-mission approach, defining the target behavior as limiting smartphone use to less than 10 minutes within each one-hour timebox. This design draws upon recent research on smartphone use regulation, which has demonstrated the effectiveness of system-driven goal-setting and interventions, where hourly micro-missions and proactive warning alarms facilitated goal adherence [88]. This hour-based approach is also rooted in prior research emphasizing the importance of setting specific and challenging goals to promote behavior change [64]. Empirical studies have also shown that limiting smartphone use to under 10 minutes per hour presents a goal that is both specific and sufficiently challenging to motivate a reduction in smartphone overuse [27, 59, 88].

3.1.2 Context-Based Micro-Incentives. For each hourly micro-mission, participants were informed in advance of the micro-incentive they would receive upon successful completion. For example, a participant might be offered 50 KRW (approximately 0.04 USD) for successfully limiting their smartphone use to under 10 minutes within a specific timebox. This micro-incentive approach is grounded in goal-setting theory [64], which underscores the importance of frequent feedback in facilitating behavior change–forming a feedback loop, enhancing awareness, and actively engaging users in the behavior change process. The incentive amounts were determined based on the assumption that an individual's adherence to the goal varies according to context and personal preferences. We hypothesized that the probability of micro-mission success would differ across three temporal contexts: *work*, *non-work*, and *weekend*. Consequently, we tailored incentive amounts separately for each context. The timing for each context was based on previous research [88]: the *work* context was defined as 9 AM to 6 PM on weekdays, the *non-work* context as 6 PM to 2 AM the following day on weekdays, and the *weekend* context encompassed all hours on weekends.

3.1.3 Multi-Objective Personalization. We framed the problem of personalization as balancing between two objectives: successful regulation of smartphone overuse and cost-effective scaffolding using financial incentives. Integrating cost-effectiveness as a core objective alongside behavior change is crucial for several reasons. First, health intervention providers typically operate within finite financial constraints. Consequently, they need to minimize incentive amounts while maintaining or enhancing the positive effects on health outcomes. Second, interestingly, research suggests that increasing financial incentive amounts does not necessarily lead to improved health outcomes [15]. Third, financial incentives can have the unintended side effect of undermining intrinsic motivation if not structured carefully–too low incentives may discourage user engagement, while too high incentives may attract individuals primarily motivated by external rewards rather than genuine behavior change [24]. Therefore, it is essential to determine appropriate incentive amounts and to minimize the overall incentive expenditure, ultimately targeting the smooth fading of financial incentives.

3.2 Mobile Application

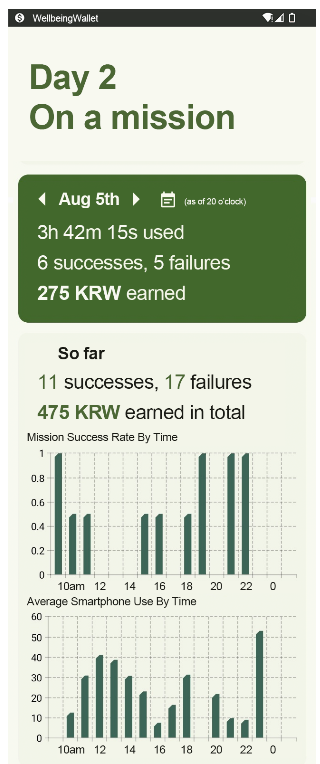
The mobile application, designed as the tool for our smartphone use regulation intervention, centered around a single core function: delivering timely notifications to users regarding their smartphone usage time and the corresponding micro-incentive amount they could earn. During the intervention period, users received hourly micro-missions, each challenging them to limit their smartphone use to under 10 minutes within a given timebox, spanning from 9 AM to 2 AM the following day. The potential incentive for successfully completing each micro-mission was determined by the participant's assigned intervention group: *personalized incentive*, *random incentive*, or *fixed incentive*. This information was readily accessible at any time through the notification bar, which displayed both the smartphone use time within the current timebox and the potential incentive for mission success. Users were able to regulate their smartphone use based on this real-time feedback. Additionally, users received notifications when their smartphone use exceeded 9 minutes and again at 10 minutes, an approach aligned with previous research where such proactive warnings enhanced self-awareness and self-monitoring for smartphone use regulation [88]. As depicted in Figure 1, users could find detailed information about their smartphone use history on the app, which included: daily smartphone usage time statistics, micro-mission success and failure counts, daily accumulated incentive amounts, overall statistics for the intervention period, and visualizations of their success rates and usage time by timebox.

During the baseline and follow-up data collection periods, no micro-missions or notifications were delivered. During the baseline period, users were informed on the app about the start date of the intervention period. During the follow-up period, users could still access their historical data on smartphone usage time and earned incentives, and the real-time smartphone use time was accessible from the notification bar. Throughout the entire study, the app collected data hourly. This data included the user ID, date, timebox, smartphone use time, micro-mission result (success or failure), and the incentive amount.

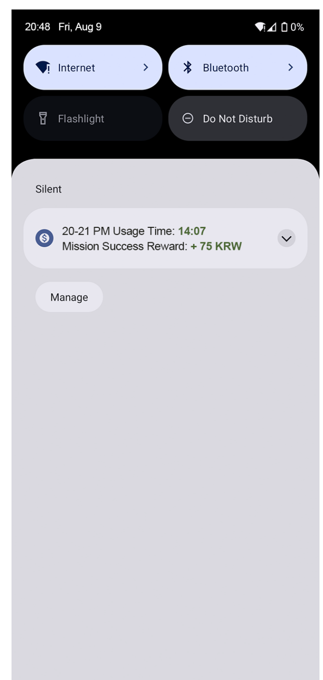
3.3 Personalization Algorithm and Control Algorithms

As presented in Algorithm 1, the personalization algorithm dynamically explored and exploited different incentive amounts to determine the optimal incentive for each user, building upon the design rationale outlined earlier–specifically, context-based micro-incentives and multi-objective personalization. The personalization problem was framed as a multi-armed bandit problem, where

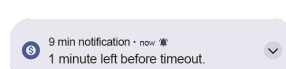
Figure 1: WellbeingWallet mobile application



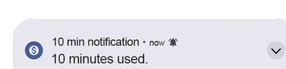
(a) Main screen of the app



(b) Notification bar timer



(c) 9-minute notification



(d) 10-minute notification

Algorithm 1: Multi-objective Thompson sampling-based incentive recommendation

Input: A set of incentive amounts, $\mathbf{I} = \{i_1, i_2, \cdots, i_K\}$

1 **Loop forever**
2   **for** $k \leftarrow 1$ **to** K **do**
3   Sample the expected success probability $\theta_k \sim \mathrm{Beta}(\alpha_k + 1, \beta_k + 1)$
4   Set the expected cost $\omega_k \leftarrow \theta_k \cdot i_k$
5   **end**
6   $k^* \leftarrow \underset{k \in [1,K]}{\arg\max}\, \theta_k \; \underset{k \in [1,K]}{\arg\min}\, \omega_k$
7   Bid $\mathbf{I}[k^*]$ for compensation of a micro-mission success
8   **if** *the micro-mission is succeeded* **then**
9     $\alpha_{k^*} \leftarrow \alpha_{k^*} + 1$
10   **else**
11     $\beta_{k^*} \leftarrow \beta_{k^*} + 1$
12   **end**
13 **end**

each arm represented the probability of successfully completing a micro-mission for that specific incentive amount and context, estimated using Thompson Sampling [101]. Three temporal contexts of *work*, *non-work*, and *weekend* and five incentive amount options of 0, 25, 50, 75, and 100 KRW (approximately 0, 0.02, 0.04, 0.06, and 0.08 USD, respectively) were used. This framework allowed comparing the probabilities of micro-mission success for each incentive amount within a given context. The prior probability distributions were initialized to 0 for the α and β values of the beta distribution, representing the counts of successes and failures of micro-missions, respectively. These parameters were updated each time a new micro-mission result was observed: if the user succeeded in the micro-mission, the α value (representing success) was incremented by one; if the user failed, the β value (representing failure) was incremented by one.

Furthermore, the personalization algorithm weighed the trade-off between maximizing behavior change success and minimizing costs, framing the problem as a multi-objective optimization task (i.e., finding the Pareto front given the two objectives, as indicated in Line 6 of Algorithm 1). For each specific context, each incentive amount option was evaluated for dominance and comparability against all other incentive options, considering both the expected success rate and expected cost. An incentive amount option was included in the Pareto front set if it was not dominated by and was comparable to the other options. From the set of Pareto-optimal options, the final incentive amount to be offered was randomly selected.

To evaluate the effectiveness of the proposed personalization algorithm, two control algorithms were used for comparison. Participants in the *fixed incentive* group consistently received a fixed incentive of 50 KRW per timebox, regardless of context or their individual performance. Participants in the *random incentive* group received a randomly selected incentive from the set of 0, 25, 50, 75, and 100 KRW per timebox, also independent of context or individual performance. Across all three intervention groups, the incentive amounts were designed to have the same expected cost of 50 KRW per timebox. Additionally, we conducted a simulation study to examine the potential of our personalization algorithm relative to the control algorithms. This simulation used hypothetical user behaviors, assuming that users would be more likely to succeed in micro-missions as incentive amounts increased. The results indicated that our personalization algorithm can encourage successful micro-missions in a cost-effective manner. Further details of the simulation study are presented in Appendix A.

4 User Study Setup

4.1 Participants

Our field study involved 72 participants recruited from online campus communities. The minimum sample size, 57, was determined using G*Power [28] for a mixed ANOVA with the following parameters [9, 17]: effect size f = 0.25 (medium), α = 0.05, power = 0.80, number of groups = 3, number of measurements = 3, correlation among repeated measures = 0.5, and nonsphericity correction ϵ = 0.5. Considering the study's long duration, a dropout rate of 20–25% was assumed. Participants were required to use Android smartphones running Android version 10.0 or higher. We also specifically sought participants who were not currently actively regulating their smartphone usage but had an intention to do so. This corresponded to individuals in either the contemplation stage (i.e., aware of the need to regulate their smartphone usage) or the preparation stage (i.e., planning to regulate their smartphone usage) of the transtheoretical model of behavior change [95]. Recognizing that prior work has shown that gender and socioeconomic status (SES) can influence perceptions of financial incentives [21, 39], we collected demographic information, including monthly spending, to facilitate stratified random sampling for our randomized controlled experiment. Given the lack of consensus on operationalizing SES for students [61, 92], we used monthly spending as a proxy for SES, based on prior research linking monthly allowance and SES within the Korean context [109].

The participants were aged between 18 and 55 (M = 27.5; SD = 7.0). Forty-nine participants were males, 22 were females, and 1 chose not to disclose their gender. Forty-one participants (57%) were high spenders, defined as spending more than 600,000 KRW (approximately 480 USD) per month based on their self-reported monthly spending. Using gender and monthly spending information, we conducted stratified random sampling [72] to distribute participants equally across three experimental groups:

- *Fixed incentive*: Participants received 50 KRW (approximately 0.04 USD[1]) for every micro-mission success.
- *Random incentive*: Participants received a randomly selected reward from the set of 0, 25, 50, 75, or 100 KRW (approximately 0, 0.02, 0.04, 0.06, 0.08 USD) for each micro-mission success.
- *Personalized incentive*: Rewards were determined using Algorithm 1, which selected one of 0, 25, 50, 75, or 100 KRW for each micro-mission success.

After stratified random sampling, the *fixed incentive* group included 23 participants (male = 70%, high spender = 57%), the *random*

[1] Note that this micro-incentive was very small, considering the average monthly household income of 4,592,321 KRW (approximately 3,674 USD) in South Korea in 2024 [60].

incentive group included 25 participants (male = 68%, high spender = 60%), and the *personalized incentive* group included 24 participants (male = 67%, high spender = 54%).

4.2 Study Procedures

Our four-week field study began with a pre-survey designed to assess participants' motivational factors related to smartphone use behavior change. This survey incorporated 37 questions from the Intrinsic Motivation Inventory (IMI) [115] and 15 questions from the Self-Regulation Questionnaire (SRQ) [116]. These theoretically grounded (i.e., self-determination theory) [81, 104] and validated [62, 73] questionnaires were employed to identify potential motivational shifts before and after the micro-interventions, such as changes in dimensions of intrinsic motivation or shifts in the locus of motivation. After completing the pre-survey, participants received instructions on installing our intervention app on their smartphones, along with essential information about the app and the micro-missions. Participants were informed that their compensation would be contingent on their mission performance. However, as this was a single-blind experiment, specific details about group differences and the underlying algorithm mechanisms were not disclosed.

The four-week study was structured as follows: The first week served as a *baseline period* for collecting smartphone usage data without any intervention. During this baseline period, our intervention app passively recorded the duration of smartphone use in the background but did not deliver any micro-missions. The second and third weeks constituted the *intervention periods*, during which our intervention app actively delivered micro-missions and provided the corresponding incentive compensation based on the assigned group. At the end of the third week, participants completed a post-survey, which included the same IMI and SRQ questionnaires as the pre-survey, as well as the Reinforcement Sensitivity Theory of Personality Questionnaire (RST-PQ) [20]. The RST-PQ, grounded in the neuropsychological theory (i.e., reinforcement sensitivity theory) [19], was used to explore potential relationships between reward sensitivity and motivation, similar to a prior study involving crowd workers [8]. The fourth and final week served as a *follow-up period*, during which all micro-financial incentives were removed. However, participants retained access to the timer feature, where they could monitor their time spent within each timebox, and to the app's main screen, where they could review their daily and overall statistics. This follow-up period was designed to investigate the sustained effects of the different intervention strategies on smartphone usage even after the removal of financial incentives. After the conclusion of the four-week field study, we conducted exit interviews with participants who volunteered to participate. Participants received a base compensation of 40,000 KRW (approximately 32 USD) for their participation and an additional compensation based on their performance in regulating their smartphone use, which ranged from a minimum of 1,075 KRW (approximately 0.86 USD) to a maximum of 12,175 KRW (approximately 9.74 USD). Participants who completed the exit interview received an additional compensation of 10,000 KRW (approximately 8 USD). This study received approval from our university's Institutional Review Board (IRB) and adhered to all the IRB guidelines (KH2024-031).

4.3 Data Collection and Analysis

The goal of our user study was to evaluate the efficacy and user experience of the personalized micro-financial incentive strategy. To achieve this, we focused on the following research questions (RQs):

- RQ1. How effective is the personalized micro-financial incentive strategy in reducing the total costs of incentives?
- RQ2. How effective is the personalized micro-financial incentive strategy in increasing the success rates of the timeboxed missions for smartphone use regulation?
  - RQ2-1. Does the suggested strategy reduce smartphone usage time compared with the baseline period?
  - RQ2-2. Does the suggested strategy reduce smartphone usage time even after removing incentives?
- RQ3. How does the personalized micro-financial incentive strategy impact intrinsic motivation and self-regulation?
- RQ4. How does the personalized micro-financial incentive strategy impact participants' perceptions and user experience?

To address these RQs, we collected both quantitative and qualitative data. Quantitative data encompassed smartphone usage data, including the user ID, date, timebox, smartphone use time, micro-mission result (success or failure), and the incentive amount for each timebox. Additionally, we collected pre- and post-survey data, including IMI scores across six sub-dimensions (interest/enjoyment, perceived competence, effort/importance, pressure/tension, perceived choice, and value/usefulness), SRQ scores across four sub-dimensions (autonomous motivation, introjected regulation, external regulation, and amotivation), RST-PQ scores, and demographic information (gender, age, and monthly spending).

Qualitative data were gathered through semi-structured exit interviews conducted with 37 participants. During these interviews, participants were asked about their overall experience with the intervention app. This included their use of different app features, perceived efficacy of the app, understanding and opinions of micro-financial incentives and their mechanism, and any perceived behavioral and motivational changes they experienced over the four-week study period. Participants were also shown visualizations of their personal data to aid recall and facilitate comparisons with their expectations. The interviews were audio-recorded, transcribed, and analyzed using reflexive thematic analysis to identify common patterns and themes [10].

5 Results

We assessed the efficacy and perceived experience of the personalized micro-financial incentives by comparing them to those of random or fixed incentives. First, the effects of different micro-financial incentive strategies were quantitatively assessed in terms of cost-effectiveness (RQ1), behavior change outcomes (RQ2), and motivational changes (RQ3). Next, the user experiences of different micro-financial incentive strategies (RQ4) were qualitatively analyzed based on insights collected from exit interviews.

5.1 RQ1. Total Costs

The *personalized incentive* strategy significantly reduced the total cost of the financial incentive program as intended through the

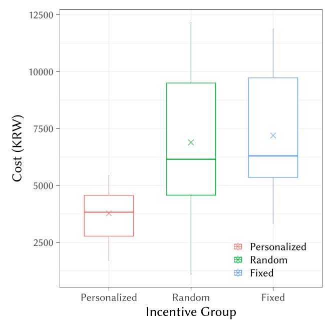


Figure 2: Average cost by incentive group

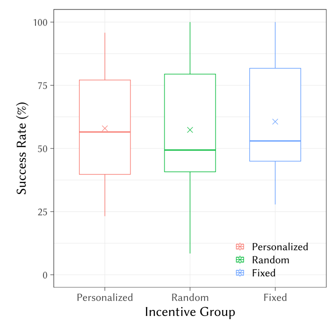


Figure 3: Average success rate by incentive group

algorithm design. The *personalized incentive* group incurred an average cost of 3,773.96 KRW (approximately 3.02 USD) ($SD = 1,102.08$; $N = 24$), the *random* group incurred 6,891.00 KRW (approximately 5.51 USD) ($SD = 3,291.79$; $N = 25$), and the *fixed* group incurred 7,193.48 KRW (approximately 5.75 USD) ($SD = 2,602.63$; $N = 23$). Figure 2 illustrates the distribution of the average cost per participant for each incentive group. The Shapiro–Wilk test confirmed the normal distribution of the total cost data for all three groups, but Levene's test indicated that the homogeneity of variance assumption was violated. Therefore, a Welch's ANOVA was performed. The results revealed a statistically significant difference in total costs between the groups, $F(2, 37.68) = 23.70$, $p < .05$. Post-hoc analysis using the Games–Howell test indicated that the total costs differed significantly between the *personalized* and *random* groups and between the *personalized* and *fixed* groups; however, there was no statistically significant difference between the *random* and *fixed* groups (Table 1).

5.2 RQ2. Behavioral Change

Figure 4 illustrates the trends in smartphone use time, measured in seconds per hour, across the four-week study period for each of the three incentive groups. As shown in Figure 3, the *personalized incentive* group achieved an average success rate of 58% ($SD = 0.21$), the *random* group 57% ($SD = 0.26$), and the *fixed* group 61% ($SD = 0.22$). Regarding the reduced smartphone use time during the intervention period, the *personalized incentive* group showed an average reduction of 339.2 seconds ($SD = 573.4$), the *random* group 412.7 seconds ($SD = 534.2$), and the *fixed* group 379.4 seconds ($SD = 443.8$). Furthermore, regarding the sustained effects of the incentive strategies, we found that during the follow-up period (compared to the baseline), the *personalized incentive* group showed an average reduction of 181.1 seconds per hour ($SD = 506.3$), the *random* group 177.9 seconds ($SD = 466.5$), and the *fixed* group 82.8 seconds ($SD = 283.7$).

To assess the effects of the different incentive strategies over time (i.e., baseline week (without interventions), intervention weeks, and follow-up week (without interventions)) on the success rate of regulating smartphone use (i.e., using less than 10 minutes per hour), a mixed ANOVA was conducted. It is important to note that neither explicit missions nor notifications were delivered during the baseline and follow-up periods. Success rates were calculated based on the log data, with success indicated if the smartphone use time was less than 10 minutes per hour. Since both normality and homoscedasticity were violated according to the Shapiro–Wilk test and the Levene's test, a robust mixed ANOVA analysis using 20% trimmed means was employed [69]. The results, detailed in Table 2, revealed a statistically significant main effect of time period on success rate. However, there was no significant main effect of incentive strategy and no significant interaction effect between incentive strategy and time period. Despite the crossovers between the *personalized* and *fixed* groups in the interaction plot in Figure 5, this interaction was not statistically significant. Post-hoc analysis using 20% trimmed means and a Bonferroni correction revealed that success rates differed significantly between the baseline and intervention periods by approximately 19%p and between the intervention and follow-up periods by approximately 16%p. However, there was no significant difference between the baseline and follow-up periods (Table 3).

5.3 RQ3. Motivational Change

Figures 6 and 7 display the changes in participants' motivation toward smartphone use regulation from pre- to post-program participation, as measured by the IMI and SRQ questionnaires. Figure 6 presents the six sub-dimensions of the IMI [115] used to

Table 1: Post-hoc analysis results for cost by incentive group

Source of variation	Mean difference ± CI	p	
Personalized - Random	3117.04 ± 1716.63	0.00	***
Personalized - Fixed	3419.52 ± 1449.85	0.00	***
Random - Fixed	302.48 ± 2067.79	0.93	

*: $p < .05$, **: $p < .01$, ***: $p < .001$; CI = 95% confidence interval

Table 2: Robust mixed ANOVA results for success rate by incentive group and intervention period

Source of variation	df	F	p	
Incentive group	2, 27.93	0.23	0.80	
Intervention period	2, 35.44	17.15	0.00	***
Incentive group × Intervention period	4, 27.53	1.14	0.36	

*: $p < .05$, **: $p < .01$, ***: $p < .001$

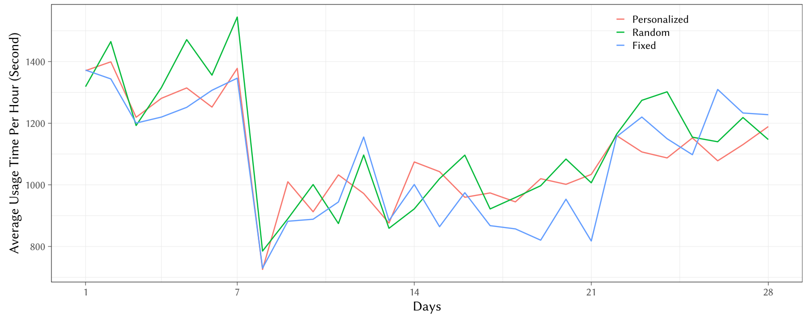


Figure 4: Average smartphone usage time per hour by incentive group

track changes in intrinsic motivation: (a) interest/enjoyment, considered the single metric for self-reported intrinsic motivation; (b) perceived competence and (e) perceived choice, both positive predictors of intrinsic motivation; (d) pressure/tension, a negative predictor of intrinsic motivation; and (c) effort/importance and (f) value/usefulness, which have been studied in relation to the continuum of motivation and regulatory styles [104]. Analysis of the IMI sub-dimension score changes revealed that the mean changes generally ranged between -1 and +1. Furthermore, **no statistically significant differences** were found in any of the IMI sub-dimensions among the three incentive groups, as determined by either a one-way ANOVA or a Kruskal–Wallis ANOVA (for non-normal distributions–i.e., value/usefulness) (Table 4). Further analysis using correlation tests showed no significant relationships between RST-PQ scores and changes in IMI sub-dimension scores.

Figure 5: Interaction between incentive group and intervention period

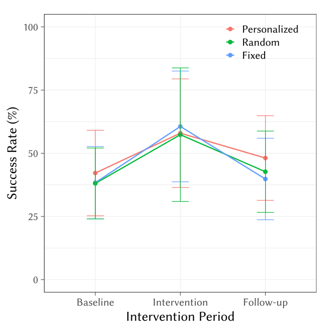


Table 3: Post-hoc analysis results for success rate by intervention period

Source of variation	Trimmed mean difference ± CI	p	
Baseline - Intervention	-0.19 ± 0.09	0.00	***
Baseline - Follow-up	-0.03 ± 0.07	1.00	
Intervention - Follow-up	0.16 ± 0.10	0.00	***

*: $p < .05$, **: $p < .01$, ***: $p < .001$; CI = 95% confidence interval

Figure 7 presents the four sub-dimensions from the SRQ [116], which focuses on changes in the locus of motivation (i.e., perceived locus of causality [102]). These sub-dimensions include: (a) autonomous motivation, closely related to self-initiated, intrinsic motivation; (b) introjected regulation and (c) external regulation, both linked to extrinsic motivation but differing in their degree of internalization; and (d) amotivation, known to be unrelated to either intrinsic or extrinsic motivation. The mean changes in SRQ sub-dimension scores generally ranged between -1 and +1. Similar to the IMI results, **no statistically significant differences** were found in any of the SRQ sub-dimensions among the three incentive groups, as confirmed using either a one-way ANOVA or a Kruskal–Wallis ANOVA (for non-normal distributions–i.e., introjected regulation and amotivation) (Table 4). Correlation analysis further indicated no significant relationships between RST-PQ scores and changes in any SRQ sub-dimension scores.

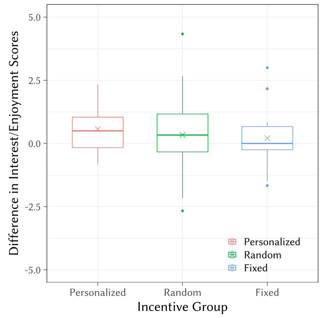
5.4 RQ4. Perceived Experience

To gain a deeper, more nuanced understanding of the user experience with the three different micro-financial incentive strategies, we conducted semi-structured interviews. We asked participants how they used the app, how app features supported their smartphone use regulation, how they understood and reacted to the

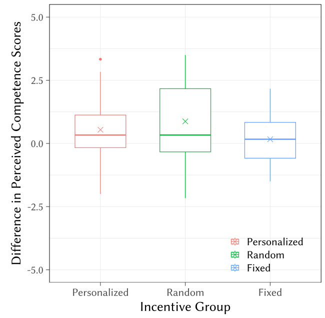
Table 4: ANOVA results for motivation change by incentive group

Source of variation	df	F	p
IMI			
Interest/enjoyment	2, 69	0.52	0.60
Perceived competence	2, 69	1.86	0.16
Effort/importance	2, 69	0.08	0.93
Pressure/tension	2, 69	1.41	0.25
Perceived choice	2, 69	0.11	0.89
Value/usefulness*	2	0.66	0.72
SRQ			
Autonomous motivation	2, 69	0.27	0.76
Introjected regulation*	2	2.29	0.32
External regulation	2, 69	0.78	0.46
Amotivation*	2	0.91	0.63

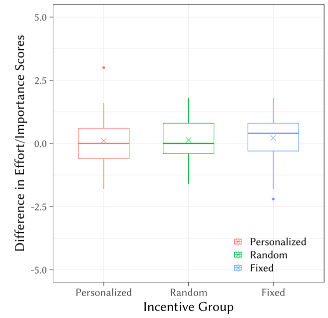
* Kruskal–Wallis ANOVA conducted



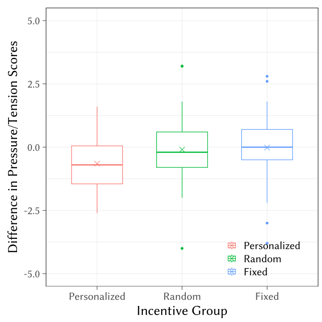
(a) Interest/enjoyment



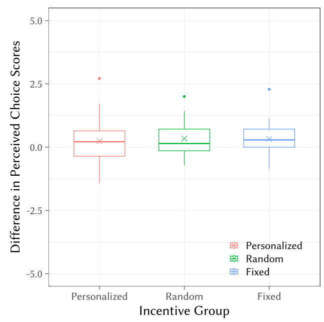
(b) Perceived competence



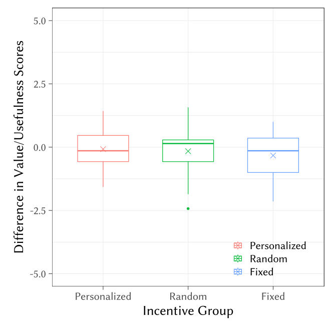
(c) Effort/importance



(d) Pressure/tension

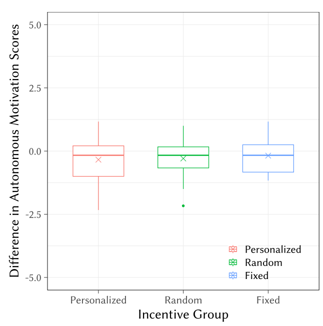


(e) Perceived choice

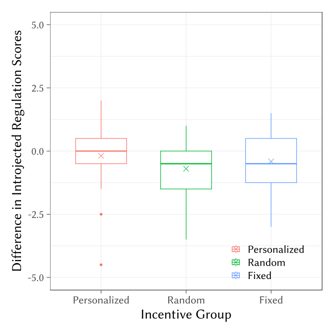


(f) Value/usefulness

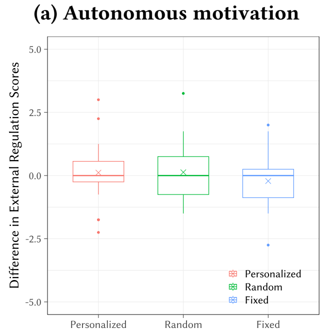
Figure 6: Change in Intrinsic Motivation Inventory (IMI) scores after intervention



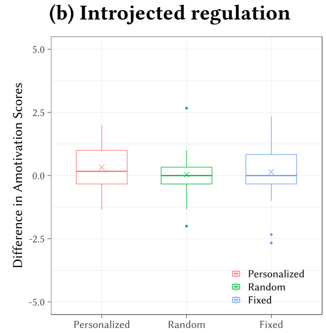
(a) Autonomous motivation



(b) Introjected regulation



(c) External regulation



(d) Amotivation

Figure 7: Change in Self-Regulation Questionnaires (SRQ) scores after intervention

micro-financial incentive mechanism, and how they perceived the changes in their behavior and motivation across the four-week study period. Participants were also shown visualizations of their data to facilitate recall and comparison with their expectations. The results are presented using the following notations for each incentive group: *P* for participants in the *personalized incentive* group, *R* for those in the *random incentive* group, and *F* for those in the *fixed incentive* group. Note that this analysis is based on the responses of the 37 interview participants.

5.4.1 Reaction to Micro-Financial Incentives. Seventeen of the 37 interview participants identified the timebox-based 9-minute warning notification as their favorite feature of the app. However, participants' responses to receiving the 9-minute warning notification varied depending on their perceptions of the financial incentives.

One group of participants ($N = 23$) reported that they did not care about the monetary incentives they would receive. In some cases, participants did not care about the small incentive amounts regardless of their assigned group. Other participants in the personalized or random incentive groups recognized the subtle fluctuations in incentive amounts but did not place much weight on the changing amounts. For instance, R05 stated, *"I was roughly aware of how much I'd failed on the day and how much I'd failed in total. So I looked at money less, and I rather looked at time spent more."* Participants in the fixed incentive group were already aware of the constant amount of incentives. Expecting the constant amount, F08 pre-planned smartphone usage: *"I use Duolingo every day. Since it normally takes 15 minutes, I planned to divide the time and complete it in two hours."*

Another group ($N = 14$) reported that they were more attentive to the incentive amounts, actively checking how much monetary incentives they would receive. They wanted to decide whether to continue or discontinue their current smartphone activity. P14, who found the monetary rewards to be the most effective aspect of the intervention, explained *"I would scroll down to check how much money I can earn in a few minutes. I checked the amount again when the 9-minute notification popped up."* Participants like P10 were more responsive to the changing amounts, reporting *"I worked hard for 100 KRW, and didn't even look at the rest. Money was definitely a motivator for me."* F23, who received the fixed amount of incentives, expressed a desire for variable rewards, suggesting, *"If there were differentials in compensations, especially for nights, I think I would have slept a little earlier."*

Notably, many participants ($N = 23$) expressed that their responses depended on the tasks that they were engaged in at the time of 9-minute notification delivery. P04 remarked, *"I would continue to use a smartphone if it's essential or beneficial to my quality of life. Reading a book, traveling by car, turning on the navigation, etc."* This suggests that participants did not solely react to the incentive amount, but rather engaged in a subjective assessment of the ongoing activity against the offered incentive amount. When asked if a different incentive amount would have altered their decision to continue or discontinue smartphone use, a majority ($N = 19$) affirmed that it would have. However, their preferences for specific incentive amounts varied according to individuals' baseline acceptance of micro-financial incentives. For instance, P10 remarked, *"If the maximum reward were 50 KRW, I would have reacted to 50 KRW,"* while R11 stated, *"If the amounts had been 10 times as much as they were, I would have been punctual and completed every mission."* Another participant, P20, expressed a preference for a predictable, fixed amount, stating, *"The amount kept changing, but it was hard to see the correlation between mission success and reward, which made it less interesting. I would have preferred a constant amount. Even after learning about the algorithm, I still can't predict how much I will earn next time. The algorithm doesn't provide enough certainty."*

Furthermore, some participants ($N = 5$) attributed their behavior more to their initial mindset than to the intervention or incentive amount. For example, R13 stated, *"I don't think it was that helpful. I didn't start with a lot of willpower, and the app itself didn't build it. I'm living a slacker life these days,"* while another participant, F07, noted, *"The reward was like a gentle nudge for those who are already motivated but hesitant to take action. Like adding a small weight to a scale that's just about to tip."* This suggests the influence of individual motivation and readiness for change on the effectiveness of micro-financial incentives.

5.4.2 Mental Model for Personalization Algorithm. Participants in the personalized or random incentive groups who noticed the fluctuations in incentive amounts reported they were curious about the underlying mechanism. For instance, R21, R22, and P10 speculated that *"more rewards were offered after spending lot of time on smartphone in the previous hour, to discourage continued smartphone use," "more rewards may have accumulated if I regulated smartphone use during the time when people normally use smartphone more, for example, from 10 PM to midnight,"* and *"rewards may have been determined based on my data collected from the first week."* Interestingly, participants assumed that the monetary incentives were designed to encourage difficult-to-achieve behaviors rather than likely-to-achieve behaviors. In other words, while the algorithm was designed to offer an incentive amount with the highest probability of success, participants expected the algorithm to suggest a higher amount for the lowest probability of success. Despite their curiosity, participants reported that they did not actively experiment with different strategies to understand or audit the algorithm. However, P14 offered an intriguing perspective regarding manipulating the algorithm: *"If I had known the mechanism of the rewards beforehand, I would have used a strategy–increasing my success rate on high bets, decreasing my success rate on low bets, and then using it after the high bets were set with high probabilities of occurrences. If I knew the algorithm, I could have continued my behavior for a longer time."* This suggests that greater transparency about the underlying algorithm could have influenced how participants engaged with the incentive system.

5.4.3 Perceived Behavioral Change. Participants reflected on their perceived changes in behavior and motivation, commenting on the visualized data of their smartphone use time and motivation scores throughout the four-week period. Participants with high mission success rates (e.g., exceeding 90%) exhibited the formation of a positive confirmatory cycle from the initial stage. R03 stated, *"The monetary reward for success is what kept me motivated. Every morning, I pressed the app, saw the graph, and felt like I lived a successful life. I often checked the report, and I felt I could do it."* On the other hand, P25 reported feeling discouraged by early failures, stating *"I realized the time was counted based on the screen being on,* *regardless of my effort. It should foster a sense of self-efficacy, where success experiences accumulate and create a positive feedback loop, but that didn't happen for me."*

Overall, the computerized setting of micro-tracking and micro-intervention demonstrated the potential of establishing a quick feedback loop, which helps to raise awareness, prompts corresponding actions, and potentially bolsters perceived competence.

When reflecting on broader behavioral changes across the entire study period, 20 participants acknowledged that their smartphone use might have increased again after the removal of the intervention. However, two participants, P09 and R15, reported subtle, yet persistent positive changes even after the intervention ended: *"Even though the program is over, I do sometimes guess if 10 minutes have already passed,"* and *"As I started using my smartphone less to earn money, I realized that I often look at it out of habit. It made me reflect. I even uninstalled the game I used to play."*

5.4.4 Perceived Motivational Change. Discussion about motivational changes, beyond observed behavioral shifts, revealed more nuanced perceptions. For example, P09 attributed their difficulties in regulating smartphone use to a decrease in their perceived competence score on the IMI, stating, *"I made an intentional effort toward smartphone use regulation, which could have resulted in the decrease in the motivation score as I realized its difficulty;"* whereas R11 interpreted similar challenges as contributing to an increase in their perceived competence score, explaining, *"The perceived competence score has increased from 2 to 5. I think it's because, even though the program has helped me systematically, I realized that I can control my time if there is external help. It's a surprising result that even such a short experience can make me feel more confident."* Similarly, P09 and P10 interpreted their decreased IMI scores for value/usefulness as an indication of a re-evaluation or a deeper, more critical understanding of smartphone use regulation, rather than as a sign of discouragement. R15 echoed this sentiment, stating, *"I realized that using a smartphone isn't necessarily a bad thing. I think my values have changed, going beyond simply increasing or decreasing."* These findings suggest that it is necessary to revisit our approach to quantifying motivation as a single flat score and to explore the multi-faceted nature of motivation in order to design more effective, tailored behavior change interventions.

6 Discussion

6.1 Role of Micro-Financial Incentives in Reinforcing Awareness

Our field study results revealed the diverse roles that micro-financial incentives played for different individuals in regulating smartphone use. We identified two primary groups: 38% of participants found the micro-financial incentives to be a motivator for digital wellbeing, whereas 62% considered them supplementary and did not pay them significant attention. Among the group highly interested in micro-financial incentives, one subgroup (71%) reported frequently checking the incentive amounts and feeling more motivated by higher amounts. In contrast, the other subgroup (29%) acknowledged their interest in and occasional checking of the incentives but maintained that they did not change their behavior or plans due to the incentive amounts. Nevertheless, they recognized that the *presence* of monetary incentives enhanced their engagement in smartphone use regulation and speculated that they would have been less involved without such additional compensation. This finding aligns with those of prior research, which demonstrated that financial incentives were not considered as the primary reason for behavior change but as an *"added bonus"* for individuals already intending to modify their health behaviors [51, 70, 96]. Mantzari et al. [70] also speculated that money might not have been the primary reason for behavior change due to factors such as the amount being too small, the influence of financial incentives operating outside participants' conscious awareness, or participants recognizing the influence of financial incentives but not acknowledging it. In this context, micro-financial incentives, as an immediate, albeit small, benefit, may have supported behavior aligned with existing motivation and facilitated self-control without drawing conscious attention to the monetary values. We posit that this characteristic of micro-financial incentives acting as an added bonus facilitated cost optimization by identifying the smallest yet still effective incentive amount for each individual.

Furthermore, micro-financial incentives demonstrated the potential to initiate a positive feedback loop, particularly when participants maintained high mission success rates in the initial stages of the intervention. The tangible accumulation of micro-financial incentives strengthened their self-efficacy, which in turn encouraged the pursuit of further financial incentives. This finding aligns with prior research on addiction treatment, where initial experiences of accomplishment and attainment have been emphasized as predictors of long-term success in behavior change [96, 117]. Conversely, negative impressions could arise if participants struggled to achieve success in the initial stages. The micro-financial incentives were designed to be provided as additional rewards for successes–i.e., in a gain frame as opposed to a loss frame [1, 88, 90]; however, a few participants interpreted their failures as losses of what they had already earned, perceiving this as unfair given the involvement of material incentives. Therefore, considering the varying roles that micro-financial incentives play for different individuals and across time, it is crucial to carefully design a micro-financial incentive strategy that adapts to the diverse and evolving needs of individuals.

6.2 Effectiveness of Multi-Armed Bandit-Based Personalization Algorithm

The proposed personalization algorithm, which leverages both multi-armed bandit and multi-objective optimization techniques, achieved comparable behavioral change outcomes to the control groups while demonstrating substantially lower costs. This was accomplished through the dynamic adjustment of incentive amounts and adaptation to individual user contexts. Unlike the fixed incentive strategy, dynamically adjusting the incentive amounts mitigated the growing insensitivity to repeated incentives over time, stimulated curiosity about upcoming incentives, and fostered continued engagement. These findings align with prior research highlighting the benefits of varied interventions [1, 48]. Unlike the purely random incentive strategy, personalizing the incentive amounts based on individual performance effectively minimized costs without compromising the intervention effectiveness or negatively affecting user experience. While both the personalized and random strategies were designed with the same initial expected cost of 50 KRW per micro-mission success, the personalization algorithm progressively adapted to individuals by exploring and exploiting different incentive amounts and accumulating user responses.

The effectiveness of the personalization algorithm suggests the potential for further enhancement, as it can be readily extended to address a variety of practical considerations beyond cost minimization. For instance, building on the core assumptions of the multi-armed bandit problem (i.e., individuals respond differently to different incentives or interventions) and the Thompson sampling method (i.e., the probabilities of behavior change success can be estimated by observing behavior and updating hypothesis distributions), diverse sets of incentive amounts can be implemented for different contexts. Furthermore, within the framework of the multi-objective optimization problem (i.e., diverse sets of objectives can be compared to each other to find the comparable and non-dominated ones), additional practical constraints can be incorporated, such as minimizing intervention fatigue by avoiding the repeated provision of certain interventions [1] or maximizing positive smartphone use [99] by predicting the utility of specific apps. Further design possibilities are discussed in the following section.

6.3 Design Considerations for Personalized Micro-Financial Incentives

6.3.1 Implementing Personalized Micro-Financial Incentives. Personalization can be implemented across various dimensions in computerized settings. Prior research from behavioral economics [48, 111] and psychology [107] provides a theoretical foundation for personalized intervention design, drawing on numerous experiments that explore the impact of varying the schedules (i.e., varying numbers or intervals), magnitudes (i.e., varying size or intensity), and immediacy (i.e., varying delay) of external stimuli (i.e., interventions) to create environments conducive to specific behaviors [74]. More specifically, fields such as contingency management [111] and applied behavior analysis (ABA) [68], although primarily focused on addiction treatment [93] and mental and developmental disorders [77], offer practical and sophisticated techniques for varying such external stimuli–such as fading (e.g., gradually reducing external stimuli to avoid reliance on them), thinning (e.g., gradually increasing the effort required to obtain external stimuli), chaining (e.g., breaking down complex behavior into smaller units for chained effects), and generalization (e.g., increasingly expanding learned behavior to new settings). Building upon this foundation, this section explores the design space for further personalizing micro-financial incentives, considering four key aspects: (1) incentive amounts, (2) incentive forms, (3) contexts, and (4) personalization objectives.

First, the amounts of micro-incentives–i.e., arms in multi-armed bandit–can be adjusted based on different rationales. While our study experimented with linearly increasing financial incentive amounts, prior research has indicated that the relationship between incentive amount and behavior change is often non-linear and independent of each other [15]. Furthermore, studies have shown that factors such as gender or SES can influence how individuals perceive financial incentives [21, 39]. Future work can explore different incentive structures, considering these non-linear, context-dependent effects and group differences.

Second, the forms of micro-incentives themselves can be varied. While maintaining the core personalization algorithm, the arms of multi-armed bandits can represent different forms of incentives beyond financial ones. For example, ABA has developed behavior change techniques that employ a variety of external stimuli, including tokens (e.g., coins, stars, points) and verbal prompts [38]. Our study combined micro-financial incentives with notification messages, and another related study has compared the impact of financial incentives and motivational messages [14]. Future research can algorithmically implement and quantitatively compare diverse forms of micro-incentives or micro-interventions to identify the most effective mix for individuals.

Third, the contexts considered for micro-incentives can be expanded beyond fixed categories such as work, non-work, and weekend times to better reflect individual needs and preferences. Many participants in our study expressed personal, context-dependent criteria for judging good and bad smartphone use (i.e., when to continue the smartphone use or not) [66], suggesting varying levels of acceptance for smartphone use regulation missions based on these individual criteria. To support mindful smartphone use and foster self-motivation, personalization algorithms could be designed to better accommodate a user's context by identifying interruptible moments based on app usage patterns or the predicted purpose of app use (e.g., for work, self-help, entertainment, or social interaction), as highlighted in JITAI research [41, 106]. A similar adaptive, just-in-time approach has been adopted by Liao et al. [63] to promote behavior change in physical activity.

Fourth, the objectives of micro-incentives can extend beyond simply balancing behavior change benefits with financial costs. For instance, participants who experienced early success in the intervention generally maintained their engagement and reported increased confidence in their ability to change their behavior, as reflected in their IMI scores. In such cases, the personalization algorithm could be designed to prioritize boosting motivation over minimizing costs. Other potential objectives include penalizing repeated arm selection to minimize intervention fatigue [1] or maximizing positive smartphone use time [99]. Furthermore, user preferences can be incorporated as an objective, as demonstrated by MyBehavior [97], which actively involves users in the personalization process. The multi-objective optimization approach can accommodate these real-world constraints and enable the development of more sophisticated personalization strategies by refining the weighting mechanism applied to different objectives. While our study employed a simple 1:1 comparison of objectives, future research can explore optimizing the weights assigned to different objectives for each individual.

Meanwhile, it should be noted that it is crucial to acknowledge and respect user control and agency when designing personalized interventions. In our study, temporal contexts were limited to fixed categories (i.e., work, non-work, and weekend) and predetermined timings (e.g., 9 AM to 6 PM for work) without accounting for individual variations in lifestyles. Similarly, the mission was uniformly defined as limiting usage to 10 minutes per hour, irrespective of individual needs or preferences. Several interview participants expressed a desire for the ability to configure their own goals or contexts. Therefore, future work should strive to balance *system-initiated personalization* with *user-initiated personalization* to better accommodate diverse lifestyles and preferences (e.g., incorporating user-initiated goal-setting).

6.3.2 Communicating Micro-Financial Incentives. Beyond the design and implementation of the personalization algorithms themselves, careful consideration should be given to how micro-financial incentives are presented to users. This section explores the design space for communicating information about (1) the personalization algorithm and (2) the micro-financial incentives to foster participant awareness.

Participants exhibited diverse mental models regarding the mechanism of varying financial incentive amounts, irrespective of whether they were assigned to the personalized or random incentive group. Examples of these mental models included: a system that offered higher incentives for timeboxes where success was predicted to be less likely; one that offered higher incentives for timeboxes where other participants typically used their smartphones more heavily (e.g., at night); one that offered higher incentives after timeboxes with excessive smartphone use; or simply a random one. Our algorithm operated under a different principle in that it suggested incentives for behaviors deemed highly likely to succeed, aiming to encourage success. This contrasted with participants' expectations that higher incentives would be used to promote less likely behaviors. Some participants expressed a lack of confidence or certainty regarding the algorithm's mechanism. In addition, even with a better understanding of the mechanism, they would have behaved similarly. However, other participants were more intrigued by the underlying mechanism and even devised strategies they believed could influence the algorithm (i.e., manipulating the algorithm to increase the probabilities of higher incentives and then benefiting from this gaming of the algorithm). These varied responses highlight the importance of transparency in communicating the algorithm's mechanism. Reducing uncertainty and enhancing predictability can empower participants with more options for planning their behavior and foster more positive behavior change by actively engaging them in the process. As suggested by prior research, providing clear explanations about the algorithm [98], visualizing the data used [42], or providing self-experimentation features [22, 50] can contribute to a deeper understanding and potentially enhance engagement.

Another critical consideration when communicating about personalized micro-financial incentives is the varying influence of micro-financial incentives on individuals. Participants assessed the micro-financial incentives using different units and reference points–e.g., the hourly incentive, total daily accumulated amount, maximum difference in hourly or daily earnings, total expected earnings over the entire study period, comparisons to the cost of everyday items like a cup of coffee or a meal, and even comparisons to the minimum hourly wage. Therefore, determining the appropriate intervention unit and schedule for presenting these incentives to each individual is likely to be a crucial factor in effectively influencing behavior change.

6.4 Potential Benefits and Limitations of Personalized Micro-Financial Incentives in Relation to Motivation Change

While prior research using financial incentives has raised concerns about the potential for undermining intrinsic motivation, i.e., crowding-out effect [103], this effect was not observed in our study. As suggested by Promberger and Marteau [96], this might be because the target behavior in our study is related to health and wellbeing, which inherently aligns with participants' intrinsic goals, thus mitigating the risk of undermining motivation. Unlike tasks used in traditional psychology or behavioral economics experiments, often "dull," "boring," or involving a conflict of interest between parties, the behavior targeted in our study may be perceived as a meaningful and interesting activity. This aligns with interview data where participants indicated that they *"already had the motivation"* to regulate their smartphone use. Thus, they tended to view the financial incentives as a positive reinforcer or, at the very least, a maintainer of their existing motivation. Moreover, our intervention did not completely restrict smartphone use, which could have negatively impacted participants' feelings of control or competence, as suggested by Reactance Theory [2]. Instead, our intervention system provided notifications and allowed participants the freedom to work toward missions at their own discretion. This approach likely preserved participants' sense of autonomy and competence, further reducing the risk of undermining effects. Additionally, while prior work has often focused on crowd workers [8, 55], where participants evaluated financial incentives in relation to effort or cost already invested, our study recruited participants who already had an intention to regulate their smartphone use and were provided financial incentives as additional rewards for future actions aligned with the desired behavior. Considering these factors, our findings support the idea that micro-financial incentives, unlike other forms of financial incentives, may lead to a *crowding-in* effect [96]–i.e., the development of new preferences favoring the incentivized behavior. This suggests that future research should investigate the longer-term effects of micro-financial incentives–e.g., habit formation [12, 65].

Despite the absence of significant differences in measured motivational changes (i.e., IMI and SRQ scores) across the incentive groups, our interview results underscore the value of examining participants' perceived experiences to gain a more nuanced understanding of motivational dynamics. For instance, some participants interpreted an increase in their perceived competence score as a positive motivational change, indicating increased confidence in their ability to regulate smartphone use, which, in turn, became a driving force for continued behavior change. Conversely, other participants acknowledged a heightened awareness and necessity of smartphone use regulation and associated this with a decrease in their perceived competence score. They perceived this increased awareness as a challenge, and it became their motivation for continued engagement. Similarly, participants interpreted both increases and decreases in perceived value/usefulness as positive changes, highlighting the complex and multi-faceted nature of interpreting motivational factors. Therefore, IMI scores should be viewed as a multi-dimensional measure of motivational change rather than a simple, unidimensional score. This more nuanced understanding of motivational changes can be incorporated into future work as one of the objectives within the personalization framework, enabling interventions to be tailored to meet individuals' specific motivational needs.

6.5 Limitations

Several limitations should be considered when interpreting the findings of our study. First, the study did not include a control group that received no micro-financial incentives, which prevented us from definitively isolating the effects of the personalized micro-financial incentives from other aspects of the intervention, such as the timeboxing mechanism and gain-framed notifications. Therefore, the observed results should be interpreted as arising from the synergistic effects of a system that integrated both reminders (a cognitive approach) and financial incentives (a behavioral approach), rather than solely from the micro-financial incentives themselves. Second, the study sample was limited to 72 participants, all residing in Korea, which may have introduced cultural influences on monetary sensitivity [109] and may not be generalizable to other populations. Third, the scope of personalization in our study was constrained by certain simplifying assumptions–e.g., our algorithm operated under a generalized timeframe for active smartphone use, assuming a common schedule for all participants: 9 AM to 2 AM the following day for overall activity, with 9 AM to 6 PM designated as working hours, 6 PM to 2 AM as non-working hours, and the entire weekend treated as a single, unified context. These broad categorizations may have limited the adaptability and overall performance of the personalized algorithm in real-world scenarios characterized by more diverse and individualized usage patterns. Fourth, the four-week duration of the study may have been insufficient to fully capture the formation of habits or the longer-term effects of the personalized algorithm.

7 Conclusion

This study introduced a personalization strategy that dynamically adjusts the amount of micro-financial incentives to promote smartphone use regulation, with the core objectives of maximizing the benefits of behavior change while simultaneously minimizing the costs associated with micro-financial incentives. Through a four-week, in-the-wild user study involving 72 participants, we investigated the efficacy and user experience of the proposed approach. Our findings demonstrated that the personalized incentive strategy was highly cost-effective in promoting smartphone use regulation without compromising intervention effectiveness. Additionally, no significant differences in motivational change were observed between the personalized, random, and fixed incentive strategies. Qualitative analysis illuminated the diverse responses to micro-financial incentives, revealing a range of mental models reagrding the personalization algorithm, and a nuanced understanding of participants' perceived behavioral and motivational changes. This study highlights the potential role of micro-financial incentives in reinforcing awareness of behavior change, acting as an *"added bonus"* to support existing motivation. Moreover, it provides empirical insights to inform the design and implementation of future personalized behavior change interventions.

Acknowledgments

We would like to thank Jiyeon Min for her contributions to the app's back-end development and the anonymous reviewers for their valuable feedback. This research was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant and the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2022-0-00064 and No. 2022R1C1C2003865, respectively).

References

[1] Palakorn Achananuparp, Ee-Peng Lim, Vibhanshu Abhishek, and Tianjiao Yun. 2018. Eat & tell: a randomized trial of random-loss incentive to increase dietary self-tracking compliance. In *Proceedings of the 2018 International Conference on Digital Health*. Association for Computing Machinery, New York, NY, USA, 45–54.
[2] Jean Adams, Emma L Giles, Elaine McColl, and Falko F Sniehotta. 2014. Carrots, sticks and health behaviours: a framework for documenting the complexity of financial incentive interventions to change health behaviours. *Health psychology review* 8, 3 (2014), 286–295.
[3] Daniel Almirall, Inbal Nahum-Shani, Nancy E Sherwood, and Susan A Murphy. 2014. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Translational behavioral medicine* 4, 3 (2014), 260–274.
[4] Mawulolo K Ameko, Miranda L Beltzer, Lihua Cai, Mehdi Boukhechba, Bethany A Teachman, and Laura E Barnes. 2020. Offline contextual multi-armed bandits for mobile health interventions: A case study on emotion regulation. In *Proceedings of the 14th ACM Conference on Recommender Systems*. 249–258.
[5] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2013. Steering user behavior with badges. In *Proceedings of the 22nd international conference on World Wide Web*. 95–106.
[6] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*. 687–698.
[7] Ionut Andone, Konrad Blaszkiewicz, Mark Eibes, Boris Trendafilov, Christian Montag, and Alexander Markowetz. 2016. Menthal: quantifying smartphone usage. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 559–564.
[8] Andrey Barashev and Guoxin Li. 2018. Worker's reward sensitivity predicting motivation in crowdsourcing: self-approach achievement goals perspective. In *Proceedings of the 1st International Conference on Information Management and Management Science*. 180–184.
[9] James Bartlett. 2019. Introduction to sample size calculation using G* Power. *European Journal of Social Psychology* (2019).
[10] Virginia Braun and Victoria Clarke. 2021. *Thematic analysis: a practical guide*. SAGE Publications Ltd.
[11] Judy Cameron and W David Pierce. 1994. Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational research* 64, 3 (1994), 363–423.
[12] Gary Charness and Uri Gneezy. 2009. Incentives to exercise. *Econometrica* 77, 3 (2009), 909–931.
[13] Mira Cheng. 2024. *Paid to Stay Sober: Promising but Debated Therapy Comes to California*. San Francisco Chronicle. https://www.sfchronicle.com/bayarea/article/drug-user-treatment-program-18575304.php
[14] Mauro Cherubini, Gabriela Villalobos-Zuñiga, Marc-Olivier Boldi, and Riccardo Bonazzi. 2020. The unexpected downside of paying or sending messages to people to make them walk: Comparing tangible rewards and motivational messages to improve physical activity. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 2 (2020), 1–44.
[15] Woohyeok Choi and Uichin Lee. 2023. Loss-Framed Adaptive Microcontingency Management for Preventing Prolonged Sedentariness: Development and Feasibility Study. *JMIR mHealth and uHealth* 11 (2023), e41660.
[16] Woohyeok Choi, Sangkeun Park, Duyeon Kim, Youn-kyung Lim, and Uichin Lee. 2019. Multi-stage receptivity model for mobile just-in-time health intervention. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 2 (2019), 1–26.
[17] Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences*. routledge.
[18] Theodore M Collins, David A Mott, Wayne E Bigelow, and David R Zimmerman. 1997. A controlled letter intervention to change prescribing behavior: results of a dual-targeted approach. *Health Services Research* 32, 4 (1997), 471.
[19] P.J. Corr. 2008. *The Reinforcement Sensitivity Theory of Personality*. Cambridge University Press. https://books.google.co.kr/books?id=d9V_uf9TgVoC
[20] Philip J Corr and Andrew J Cooper. 2016. The reinforcement sensitivity theory of personality questionnaire (RST-PQ): development and validation. *Psychological*

assessment 28, 11 (2016), 1427.

[21] Rachel Croson and Uri Gneezy. 2009. Gender differences in preferences. *Journal of Economic literature* 47, 2 (2009), 448–474.

[22] Nediyana Daskalova, Jina Yoon, Yibing Wang, Cintia Araujo, Guillermo Beltran Jr, Nicole Nugent, John McGeary, Joseph Jay Williams, and Jeff Huang. 2020. SleepBandits: Guided flexible self-experiments for sleep. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[23] Edward L Deci. 1971. Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology* 18, 1 (1971), 105.

[24] Edward L Deci, Richard Koestner, and Richard M Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125, 6 (1999), 627.

[25] Anthony DeFulio and Kenneth Silverman. 2012. The use of incentives to reinforce medication adherence. *Preventive medicine* 55 (2012), S86–S94.

[26] Ryan J Dwyer, Kostadin Kushlev, and Elizabeth W Dunn. 2018. Smartphone use undermines enjoyment of face-to-face social interactions. *Journal of Experimental Social Psychology* 78 (2018), 233–239.

[27] Hossein Falaki, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. 2010. Diversity in smartphone usage. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*. 179–194.

[28] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.

[29] Andrew T Fiore, Coye Cheshire, Lindsay Shaw Taylor, and GA Mendelsohn. 2014. Incentives to participate in online research: an experimental examination of" surprise" incentives. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3433–3442.

[30] Brianna S Fjeldsoe, Alison L Marshall, and Yvette D Miller. 2009. Behavior change interventions delivered by mobile telephone short-message service. *American journal of preventive medicine* 36, 2 (2009), 165–173.

[31] National Institute for Health and Clinical Excellence. 2010. *The use of incentives to improve health: Citizens Council Meeting*. London National Institute for Health and Clinical Excellence.

[32] Evan M Forman, Stephanie P Goldstein, Rebecca J Crochiere, Meghan L Butryn, Adrienne S Juarascio, Fengqing Zhang, and Gary D Foster. 2019. Randomized controlled trial of OnTrack, a just-in-time adaptive intervention designed to enhance weight loss. *Translational behavioral medicine* 9, 6 (2019), 989–1001.

[33] Bruno S Frey et al. 1997. On the relationship between intrinsic and extrinsic work motivation. *International journal of industrial organization* 15, 4 (1997), 427–439.

[34] Bruno S Frey and Reto Jegen. 2001. Motivation crowding theory. *Journal of economic surveys* 15, 5 (2001), 589–611.

[35] Matthew Fuller-Tyszkiewicz, Ben Richardson, Vivienne Lewis, Jake Linardon, Jacqueline Mills, Kerry Juknaitis, Charlotte Lewis, Kim Coulson, Renee O'Donnell, Lilani Arulkadacham, et al. 2019. A randomized trial exploring mindfulness and gratitude exercises as eHealth-based micro-interventions for improving body satisfaction. *Computers in Human Behavior* 95 (2019), 58–65.

[36] David Gibson, Nathaniel Ostashewski, Kim Flintoff, Sheryl Grant, and Erin Knight. 2015. Digital badges in education. *Education and Information Technologies* 20 (2015), 403–410.

[37] Uri Gneezy, Stephan Meier, and Pedro Rey-Biel. 2011. When and why incentives (don't) work to modify behavior. *Journal of economic perspectives* 25, 4 (2011), 191–210.

[38] Timothy D Hackenberg. 2009. Token reinforcement: A review and analysis. *Journal of the experimental analysis of behavior* 91, 2 (2009), 257–286.

[39] Nancy Haff, Mitesh S Patel, Raymond Lim, Jingsan Zhu, Andrea B Troxel, David A Asch, and Kevin G Volpp. 2015. The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *American Journal of Health Promotion* 29, 5 (2015), 314–323.

[40] Lasse Hakulinen, Tapio Auvinen, and Ari Korhonen. 2015. The effect of achievement badges on students' behavior: An empirical study in a university-level computer science course. *International Journal of Emerging Technologies in Learning* 10, 1 (2015).

[41] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. 2019. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *International Journal of Behavioral Nutrition and Physical Activity* 16 (2019), 1–21.

[42] Jeffrey Heer and Ben Shneiderman. 2012. Interactive dynamics for visual analysis: A taxonomy of tools that support the fluent and flexible use of visualizations. *Queue* 10, 2 (2012), 30–55.

[43] Kristin E Heron and Joshua M Smyth. 2010. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *British journal of health psychology* 15, 1 (2010), 1–39.

[44] Alexis Hiniker, Sungsoo Hong, Tadayoshi Kohno, and Julie A Kientz. 2016. MyTime: designing and evaluating an intervention for smartphone non-use. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 4746–4757.

[45] Esther Howe, Jina Suh, Mehrab Bin Morshed, Daniel McDuff, Kael Rowan, Javier Hernandez, Marah Ihab Abdin, Gonzalo Ramos, Tracy Tran, and Mary P Czerwinski. 2022. Design of digital workplace stress-reduction intervention systems: Effects of intervention type and timing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.

[46] Caroline Jaffe, Cristina Mata, and Sepandar Kamvar. 2017. Motivating urban cycling through a blockchain-based financial incentives system. In *Proceedings of the 2017 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2017 ACM international symposium on wearable computers*. 81–84.

[47] Robert W Jeffery. 2012. Financial incentives and weight control. *Preventive medicine* 55 (2012), S61–S67.

[48] Daniel Kahneman and Amos Tversky. 2013. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 99–127.

[49] Dominic Kao and D Fox Harrell. 2018. The effects of badges and avatar identification on play and making in educational games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–19.

[50] Ravi Karkar, Jasmine Zia, Roger Vilardaga, Sonali R Mishra, James Fogarty, Sean A Munson, and Julie A Kientz. 2016. A framework for self-experimentation in personalized health. *Journal of the American Medical Informatics Association* 23, 3 (2016), 440–448.

[51] Annice Kim, Kian Kamyab, Jingsan Zhu, and Kevin Volpp. 2011. Why are financial incentives not effective at influencing some smokers to quit? Results of a process evaluation of a worksite trial assessing the efficacy of financial incentives for smoking cessation. *Journal of occupational and environmental medicine* 53, 1 (2011), 62–67.

[52] Jihoon Kim and Darla M Castelli. 2021. Effects of gamification on behavioral change in education: A meta-analysis. *International Journal of Environmental Research and Public Health* 18, 7 (2021), 3550.

[53] Jaejeung Kim, Hayoung Jung, Minsam Ko, and Uichin Lee. 2019. Goalkeeper: Exploring interaction lockout mechanisms for regulating smartphone use. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–29.

[54] Jaejeung Kim, Joonyoung Park, Hyunsoo Lee, Minsam Ko, and Uichin Lee. 2019. LocknType: Lockout task intervention for discouraging smartphone app use. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[55] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1301–1318.

[56] Predrag Klasnja, Shawna Smith, Nicholas J Seewald, Andy Lee, Kelly Hall, Brook Luers, Eric B Hekler, and Susan A Murphy. 2019. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. *Annals of Behavioral Medicine* 53, 6 (2019), 573–582.

[57] Minsam Ko, Seungwoo Choi, Koji Yatani, and Uichin Lee. 2016. Lock n'LoL: group-based limiting assistance app to mitigate smartphone distractions in group activities. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 998–1010.

[58] Minsam Ko, Chayanin Wong, Sunmin Son, Euigon Jung, Uichin Lee, Seungwoo Choi, Sungho Jo, and Min H Kim. 2015. Lock n'LoL: Mitigating Smartphone Disturbance in Co-located Social Interactions. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. 1561–1566.

[59] Minsam Ko, Subin Yang, Joonwon Lee, Christian Heizmann, Jinyoung Jeong, Uichin Lee, Daehee Shin, Koji Yatani, Junehwa Song, and Kyong-Mee Chung. 2015. NUGU: a group-based intervention app for improving self-regulation of limiting smartphone use. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 1235–1245.

[60] Statistics Korea. 2024. *2024.3/4, All households (average), Income.* Statistics Korea. https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1L9U018&conn_path=I2&language=en

[61] Annette Lareau and Dalton Conley. 2008. *Social class: How does it work?* Russell Sage Foundation.

[62] Chantal S Levesque, Geoffrey C Williams, Diane Elliot, Michael A Pickering, Bradley Bodenhamer, and Phillip J Finley. 2007. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. *Health education research* 22, 5 (2007), 691–702.

[63] Peng Liao, Kristjan Greenewald, Predrag Klasnja, and Susan Murphy. 2020. Personalized heartsteps: A reinforcement learning algorithm for optimizing physical activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–22.

[64] Edwin A Locke and Gary P Latham. 2006. New directions in goal-setting theory. *Current directions in psychological science* 15, 5 (2006), 265–268.

[65] George Loewenstein, Joseph Price, and Kevin Volpp. 2016. Habit formation in children: Evidence from incentives for healthy eating. *Journal of health*

economics 45 (2016), 47–54.
[66] Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. 2018. What makes smartphone use meaningful or meaningless? *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–26.
[67] Jennifer Plebani Lussier, Sarah H Heil, Joan A Mongeon, Gary J Badger, and Stephen T Higgins. 2006. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction* 101, 2 (2006), 192–203.
[68] Gregory J Madden, William V Dube, Timothy D Hackenberg, Gregory P Hanley, and Kennon A Lattal. 2013. *APA handbook of behavior analysis, Vol. 1: Methods and principles.* American Psychological Association.
[69] Patrick Mair and Rand Wilcox. 2020. Robust statistical methods in R using the WRS2 package. *Behavior research methods* 52 (2020), 464–488.
[70] Eleni Mantzari, Florian Vogt, and Theresa M Marteau. 2012. The effectiveness of financial incentives for smoking cessation during pregnancy: is it from being paid or from the extra aid? *BMC Pregnancy and Childbirth* 12 (2012), 1–12.
[71] Eleni Mantzari, Florian Vogt, Ian Shemilt, Yinghui Wei, Julian PT Higgins, and Theresa M Marteau. 2015. Personal financial incentives for changing habitual health-related behaviors: a systematic review and meta-analysis. *Preventive medicine* 75 (2015), 75–85.
[72] Jeovany Martínez-Mesa, David Alejandro González-Chica, Rodrigo Pereira Duquia, Renan Rangel Bonamigo, and João Luiz Bastos. 2016. Sampling: how to select participants in my research study? *Anais brasileiros de dermatologia* 91, 3 (2016), 326–330.
[73] Edward McAuley, Terry Duncan, and Vance V Tammen. 1989. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport* 60, 1 (1989), 48–58.
[74] Sterling M McPherson, Sara Parent, André Miguel, and Michael McDonell. 2022. Contingency Management Is a Powerful Clinical Tool for Treating Substance Use: Research Evidence and New Practice Guidelines for Use. *Psychiatric Times* 39, 9 (2022).
[75] Gunther Meinlschmidt, Jong-Hwan Lee, Esther Stalujanis, Angelo Belardi, Minkyung Oh, Eun Kyung Jung, Hyun-Chul Kim, Janine Alfano, Seung-Schik Yoo, and Marion Tegethoff. 2016. Smartphone-based psychotherapeutic micro-interventions to improve mood in a real-world setting. *Frontiers in psychology* 7 (2016), 1112.
[76] Carla K Miller. 2019. Adaptive intervention designs to promote behavioral change in adults: what is the evidence? *Current diabetes reports* 19 (2019), 1–9.
[77] Raymond G Miltenberger. 2016. *Behavior modification: Principles and procedures.* Cengage Learning.
[78] Marc S Mitchell, Jack M Goodman, David A Alter, Leslie K John, Paul I Oh, Maureen T Pakosh, and Guy E Faulkner. 2013. Financial incentives for exercise adherence in adults: systematic review and meta-analysis. *American journal of preventive medicine* 45, 5 (2013), 658–667.
[79] Marc S Mitchell, Stephanie L Orstad, Aviroop Biswas, Paul I Oh, Melanie Jay, Maureen T Pakosh, and Guy Faulkner. 2020. Financial incentives for physical activity in adults: systematic review and meta-analysis. *British Journal of Sports Medicine* 54, 21 (2020), 1259–1268.
[80] Alberto Monge Roffarello and Luigi De Russis. 2019. The race towards digital wellbeing: Issues and opportunities. In *Proceedings of the 2019 CHI conference on human factors in computing systems.* 1–14.
[81] Vera Monteiro, Lourdes Mata, and Francisco Peixoto. 2015. Intrinsic motivation inventory: Psychometric properties in the context of first language and mathematics learning. *Psicologia: Reflexão e Crítica* 28, 3 (2015), 434–443.
[82] Mohamed Musthag, Andrew Raij, Deepak Ganesan, Santosh Kumar, and Saul Shiffman. 2011. Exploring micro-incentive strategies for participant compensation in high-burden studies. In *Proceedings of the 13th international conference on Ubiquitous computing.* 435–444.
[83] Inbal Nahum-Shani, Eric B Hekler, and Donna Spruijt-Metz. 2015. Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health psychology* 34, S (2015), 1209.
[84] Claudia Niza, Burcu Tung, and Theresa M Marteau. 2013. Incentivizing blood donation: Systematic review and meta-analysis to test Titmuss' hypotheses. *Health Psychology* 32, 9 (2013), 941.
[85] Fabian Okeke, Michael Sobolev, Nicola Dell, and Deborah Estrin. 2018. Good vibrations: can a digital nudge reduce digital overload?. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services.* 1–12.
[86] Adiba Orzikulova, Han Xiao, Zhipeng Li, Yukang Yan, Yuntao Wang, Yuanchun Shi, Marzyeh Ghassemi, Sung-Ju Lee, Anind K Dey, and Xuhai Xu. 2024. Time2Stop: Adaptive and Explainable Human-AI Loop for Smartphone Overuse Intervention. In *Proceedings of the CHI Conference on Human Factors in Computing Systems.* 1–20.
[87] Pablo Paredes, Ran Gilad-Bachrach, Mary Czerwinski, Asta Roseway, Kael Rowan, and Javier Hernandez. 2014. PopTherapy: Coping with stress through pop-culture. In *Proceedings of the 8th international conference on pervasive computing technologies for healthcare.* 109–117.
[88] Joonyoung Park, Hyunsoo Lee, Sangkeun Park, Kyong-Mee Chung, and Uichin Lee. 2021. Goldentime: Exploring system-driven timeboxing and micro-financial incentives for self-regulated phone use. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–17.
[89] Joonyoung Park, Jin Yong Sim, Jaejeung Kim, Mun Yong Yi, and Uichin Lee. 2018. Interaction restraint: enforcing adaptive cognitive tasks to restrain problematic user interaction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–6.
[90] Mitesh S Patel, David A Asch, Roy Rosin, Dylan S Small, Scarlett L Bellamy, Jack Heuer, Susan Sproat, Chris Hyson, Nancy Haff, Samantha M Lee, et al. 2016. Framing financial incentives to increase physical activity among overweight and obese adults: a randomized, controlled trial. *Annals of internal medicine* 164, 6 (2016), 385–394.
[91] Robert Pear. 2013. *Employers Get Leeway on Health Incentives.* The New York Times. https://www.nytimes.com/2013/05/30/business/new-rules-give-employers-leeway-on-use-of-health-incentives.html
[92] Gwen Petro, Amy Gonzales, and Jessica Calarco. 2020. " Out of Luck": Socio-Economic Differences in Student Coping Responses to Technology Problems. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* 1–10.
[93] Nancy M Petry. 2000. A comprehensive guide to the application of contingency management procedures in clinical settings. *Drug and alcohol dependence* 58, 1-2 (2000), 9–25.
[94] Nancy M Petry and Francis Simcic Jr. 2002. Recent advances in the dissemination of contingency management techniques: Clinical and research perspectives. *Journal of substance abuse treatment* 23, 2 (2002), 81–86.
[95] James O Prochaska and Wayne F Velicer. 1997. The transtheoretical model of health behavior change. *American journal of health promotion* 12, 1 (1997), 38–48.
[96] Marianne Promberger and Theresa M Marteau. 2013. When do financial incentives reduce intrinsic motivation? comparing behaviors studied in psychological and economic literatures. *Health Psychology* 32, 9 (2013), 950.
[97] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, and Tanzeem Choudhury. 2015. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing.* 707–718.
[98] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. " Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* 1135–1144.
[99] Alberto Monge Roffarello and Luigi De Russis. 2023. Achieving digital wellbeing through digital self-control tools: A systematic review and meta-analysis. *ACM Transactions on Computer-Human Interaction* 30, 4 (2023), 1–66.
[100] John Rooksby, Parvin Asadzadeh, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2016. Personal tracking of screen time on digital devices. In *Proceedings of the 2016 CHI conference on human factors in computing systems.* 284–296.
[101] Daniel J Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, Zheng Wen, et al. 2018. A tutorial on thompson sampling. *Foundations and Trends® in Machine Learning* 11, 1 (2018), 1–96.
[102] Richard M Ryan and James P Connell. 1989. Perceived locus of causality and internalization: examining reasons for acting in two domains. *Journal of personality and social psychology* 57, 5 (1989), 749.
[103] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.
[104] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.
[105] Hillol Sarker, Moushumi Sharmin, Amin Ahsan Ali, Md Mahbubur Rahman, Rummana Bari, Syed Monowar Hossain, and Santosh Kumar. 2014. Assessing the availability of users to engage in just-in-time intervention in the natural environment. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing.* 909–920.
[106] Susan M Schembre, Yue Liao, Michael C Robertson, Genevieve Fridlund Dunton, Jacqueline Kerr, Meghan E Haffey, Taylor Burnett, Karen Basen-Engquist, and Rachel S Hicklen. 2018. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. *Journal of medical Internet research* 20, 3 (2018), e106.
[107] Burrhus Frederic Skinner. 1965. *Science and human behavior.* Number 92904. Simon and Schuster.
[108] Michael Sobolev, Fabian Okeke, and Ori Plonsky. 2023. mRAPID Study: Effect of Micro-incentives and Daily Deadlines on Practice Behavior. In *International Conference on Persuasive Technology.* Springer, 67–81.
[109] Sang-Hee Sohn, So-Hyun Joo, John E Grable, Seonglim Lee, and Minjeung Kim. 2012. Adolescents' financial literacy: The role of financial socialization agents, financial experiences, and money attitudes in shaping financial literacy among South Korean youth. *Journal of adolescence* 35, 4 (2012), 969–980.

[110] Stephanie A Spohr, Rajesh Nandy, Deepthi Gandhiraj, Abhilash Vemulapalli, Sruthi Anne, and Scott T Walters. 2015. Efficacy of SMS text message interventions for smoking cessation: a meta-analysis. *Journal of substance abuse treatment* 56 (2015), 1–10.

[111] John ER Staddon and Daniel T Cerutti. 2003. Operant conditioning. *Annual review of psychology* 54, 1 (2003), 115–144.

[112] Catherine Stanger, Emily A Scherer, Steven F Babbin, Stacy R Ryan, and Alan J Budney. 2017. Abstinence based incentives plus parent training for adolescent alcohol and other substance misuse. *Psychology of Addictive Behaviors* 31, 4 (2017), 385.

[113] Juho Sun, Sangkeun Park, Gyuwon Jung, Yong Jeong, Uichin Lee, Kyong-Mee Chung, Changseok Lee, Heewon Kim, Suhyon Ahn, Ahsan Khandoker, et al. 2020. BeActive: Encouraging physical activities with just-in-time health intervention and micro financial incentives. In *Proceedings of the 2020 Symposium on Emerging Research from Asia and on Asian Contexts and Cultures*. 17–20.

[114] Kim Sutherland, Jon B Christianson, and Sheila Leatherman. 2008. Impact of targeted financial incentives on personal health behavior. *Medical Care Research and Review* 65, 6_suppl (2008), 36S–78S.

[115] Center For Self-Determination Theory. 2024. Intrinsic Motivation Inventory (IMI). https://selfdeterminationtheory.org/intrinsic-motivation-inventory/

[116] Center For Self-Determination Theory. 2024. Self-Regulation Questionnaires (SRQ). https://selfdeterminationtheory.org/self-regulation-questionnaires/

[117] P Tonnesen, P Paoletti, G Gustavsson, MA Russell, R Saracci, A Gulsvik, B Rijcken, and U Sawe. 1999. Higher dosage nicotine patches increase one-year smoking cessation rates: results from the European CEASE trial. Collaborative European Anti-Smoking Evaluation. European Respiratory Society. *European Respiratory Journal* 13, 2 (1999), 238–246.

[118] Liam D Turner, Stuart M Allen, and Roger M Whitaker. 2017. Reachable but not receptive: Enhancing smartphone interruptibility prediction by modelling the extent of user engagement with notifications. *Pervasive and Mobile Computing* 40 (2017), 480–494.

[119] Zahra Vahedi and Alyssa Saiphoo. 2018. The association between smartphone use, stress, and anxiety: A meta-analytic review. *Stress and Health* 34, 3 (2018), 347–358.

[120] Rama Adithya Varanasi, Nicola Dell, and Aditya Vashistha. 2024. Saharaline: A Collective Social Support Intervention for Teachers in Low-Income Indian Schools. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[121] Ivo Vlaev, Dominic King, Ara Darzi, and Paul Dolan. 2019. Changing health behaviors using financial incentives: a review from behavioral economics. *BMC public health* 19 (2019), 1–9.

[122] Scott T Walters, Eun-Young Mun, Zhengqi Tan, Justin M Luningham, Emily T Hébert, Jason A Oliver, and Michael S Businelle. 2022. Development and preliminary effectiveness of a smartphone-based, just-in-time adaptive intervention for adults with alcohol misuse who are experiencing homelessness. *Alcoholism: Clinical and Experimental Research* 46, 9 (2022), 1732–1741.

[123] Liyuan Wang and Lynn Carol Miller. 2020. Just-in-the-moment adaptive interventions (JITAI): a meta-analytical review. *Health Communication* 35, 12 (2020), 1531–1544.

[124] Ruolan Wu, Chun Yu, Xiaole Pan, Yujia Liu, Ningning Zhang, Yue Fu, Yuhan Wang, Zhi Zheng, Li Chen, Qiaolei Jiang, et al. 2024. MindShift: Leveraging Large Language Models for Mental-States-Based Problematic Smartphone Use Intervention. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–24.

[125] Xuhai Xu, Tianyuan Zou, Han Xiao, Yanzhang Li, Ruolin Wang, Tianyi Yuan, Yuntao Wang, Yuanchun Shi, Jennifer Mankoff, and Anind K Dey. 2022. TypeOut: leveraging just-in-time self-affirmation for smartphone overuse reduction. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[126] Tetsuo Yamabe, Vili Lehdonvirta, Hitoshi Ito, Hayuru Soma, Hiroaki Kimura, and Tatsuo Nakajima. 2010. Activity-based micro-pricing: Realizing sustainable behavior changes through economic incentives. In *Persuasive Technology: 5th International Conference, PERSUASIVE 2010, Copenhagen, Denmark, June 7-10, 2010. Proceedings 5*. Springer, 193–204.

[127] Min-Jeong Yang, Steven K Sutton, Laura M Hernandez, Sarah R Jones, David W Wetter, Santosh Kumar, and Christine Vinci. 2023. A Just-In-Time Adaptive intervention (JITAI) for smoking cessation: Feasibility and acceptability findings. *Addictive behaviors* 136 (2023), 107467.

[128] Stav Yanovsky, Nicholas Hoernle, Omer Lev, and Kobi Gal. 2019. One size does not fit all: Badge behavior in q&a sites. In *Proceedings of the 27th ACM conference on user modeling, adaptation and personalization*. 113–120.

[129] Fengpeng Yuan, Xianyi Gao, and Janne Lindqvist. 2017. How busy are you? Predicting the interruptibility intensity of mobile users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 5346–5360.

[130] Rui Zhen, Ru-De Liu, Wei Hong, and Xiao Zhou. 2019. How do interpersonal relationships relieve adolescents' problematic mobile phone use? The roles of loneliness and motivation to use mobile phones. *International journal of environmental research and public health* 16, 13 (2019), 2286.

Appendix

A Simulation Study Results

This appendix presents the results of a simulation study designed to validate that our personalization algorithm effectively handles various user behaviors in response to different incentive amounts. To this end, we randomly generated hypothetical user behaviors, assuming that users are more likely to succeed in behavioral missions as the incentive amount increases. Using these generated behaviors, we conducted simulations comparing the performance of different incentive algorithms–personalized, random, and fixed–in terms of success rates, total costs, and cost-effectiveness. The results confirmed that our personalization incentive algorithm can encourage users to succeed in behavioral missions in a cost-effective manner.

A.1 Hypothetical User Behaviors

To evaluate whether our personalization algorithm could effectively handle a wide range of user behaviors toward different incentive amounts, we generated a diverse set of hypothetical user behavior profiles. Based on prior research indicating that larger incentives have a more significant effect on promoting health behavior change [67, 79], we represented each hypothetical user behavior as a monotonically increasing function, where the domain comprises different incentive amounts, and the range indicates the probability of success in behavior missions. Among various monotonic functions, we employed the logistic regression function owing to its natural ability to map incentive amounts to probabilities, as follows:

$$y = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x)]}$$

To ensure that this equation yields probabilities near zero (i.e., 0.001) and near one (i.e., 0.999) at the minimum and maximum incentive amounts, respectively, the parameters β_0 and β_1 must be set appropriately. Note that because the logistic regression function asymptotically approaches probabilities of 0.0 and 1.0 at negative and positive infinities, respectively, we considered near-zero and near-one probabilities instead of exact values. To determine β_0 and β_1, we set up the following equations:

$$0.001 = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 r_{\min})]}, \quad 0.999 = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 r_{\max})]}$$

where $r_{\min}$ and $r_{\max}$ are the minimum and maximum incentive amounts, respectively. Solving these equations yields the following:

$$\beta_0 = -\frac{2 \ln 999 (r_{\max} + r_{\min})}{r_{\max} - r_{\min}}, \quad \beta_1 = \frac{2 \ln 999}{r_{\max} - r_{\min}}$$

Thus, our logistic regression for hypothetical user behavior becomes:

$$\begin{aligned} y &= \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x)]} \\ &= \frac{1}{1 + \exp\left[-\left(-\frac{2 \ln 999 (r_{\max} + r_{\min})}{r_{\max} - r_{\min}} + \frac{2 \ln 999}{r_{\max} - r_{\min}} x\right)\right]} \\ &= \frac{1}{1 + \exp\left[-\left(2 \ln 999 \frac{x - r_{\max} - r_{\min}}{r_{\max} - r_{\min}}\right)\right]} \end{aligned}$$

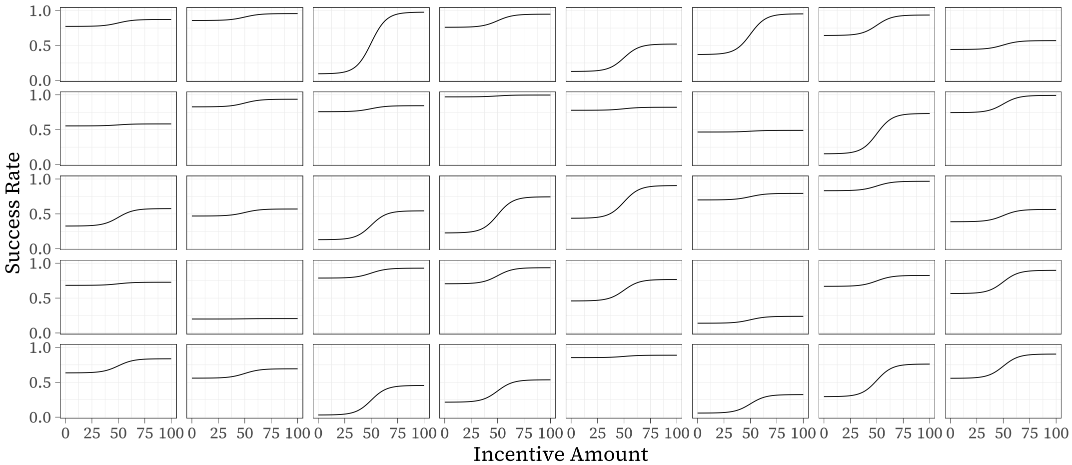


Figure 8: Examples of hypothetical user behaviors by varying parameters, $p_{\min}$ and $p_{\max}$

The above equation produces outputs ranging from 0.001 to 0.999. However, real users may exhibit behavior change probabilities greater than 0.001 at the minimum incentive amount and less than 0.999 at the maximum incentive amount owing to individual differences. Therefore, we rescaled the outputs using a specified range, $[p_{\min}, p_{\max}]$, as follows:

$$y = p_{\min} + (p_{\max} - p_{\min}) \frac{1}{1 + \exp\left[-\left(2 \ln 999 \frac{x - r_{\max} - r_{\min}}{r_{\max} - r_{\min}}\right)\right]}$$

where $p_{\min}$ and $p_{\max}$ represent the minimum and maximum probabilities of behavior change, respectively.

A.2 Procedure

Our simulation study evaluated three incentive algorithms: fixed, random, and personalized, consistent with our user study. The fixed incentive provided 50 KRW for each successful behavior mission. The random incentive randomly selected one of the following amounts: 0, 25, 50, 75, or 100 KRW for each mission. For the personalized incentive, the incentive amount was determined by our algorithm (Algorithm 1 in the main text). The minimum and maximum incentive amounts were set at 0 and 100 KRW, respectively (i.e., $r_{\min} = 0, r_{\max} = 100$). We generated 1,000 hypothetical user behaviors by systematically varying the parameters $p_{\min}$ and $p_{\max}$ (Figure 8).

Each simulation run began with the incentive algorithm suggesting a specific incentive amount. The hypothetical user behavior then returned a probability indicating the likelihood of a successful behavior mission for that incentive amount. This interaction was repeated until 500 behavior missions were successfully completed. We conducted simulations for all possible pairs of incentive algorithms and user behaviors (i.e., three algorithms × 1,000 behaviors = 3,000 pairs) and measured the success rate, total costs, and cost-per-success.

A.3 Results

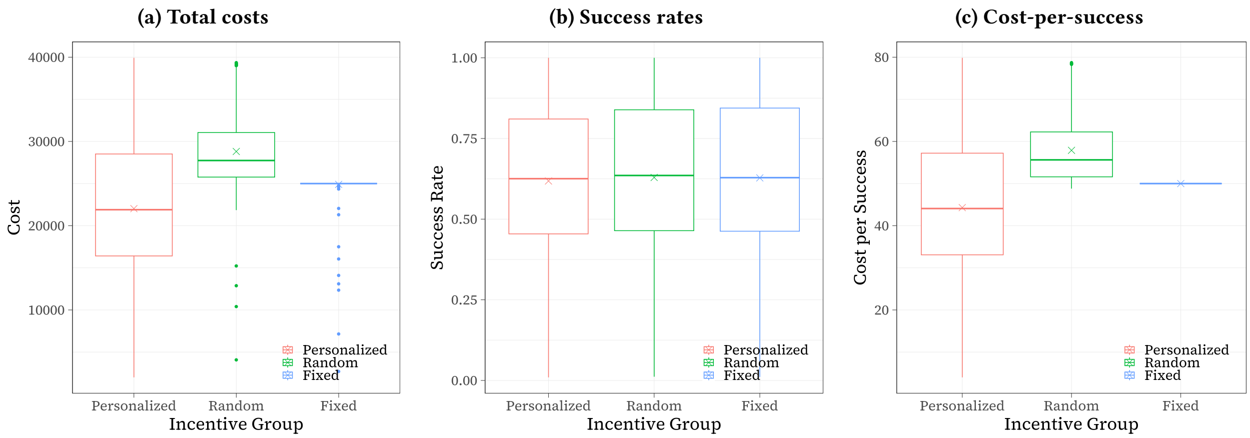
Our simulation study demonstrated that the proposed personalized incentive algorithm could elicit behavior change cost-effectively. As shown in Figure 9, the personalized incentive model incurred a total cost of 22,025.13 KRW ($SD = 8,614.61$) for 500 mission successes, which was lower than 28,808.70 KRW ($SD = 3,956.03$) for the random incentive and 24,900.00 KRW ($SD = 1,177.77$) for the fixed incentive. Success rates were similar across the three incentive algorithms: 62% ($SD = 0.24$) for the personalized incentive, 63% ($SD = 0.24$) for the random incentive, and 63% ($SD = 0.24$) for the fixed incentive. Additionally, our personalized mechanism compensated 44.30 KRW ($SD = 16.98$) per successful mission, which was less than 57.89 KRW ($SD = 7.60$) for the random incentive and 50.00 KRW ($SD = 0.00$) for the fixed incentive.

Our personalized incentive design seeks to minimize compensation costs by recommending a small incentive amount that is likely to trigger behavior change. This strategy suggests that cost-effectiveness can be maximized when users are highly likely to adjust their behavior with a minimal incentive. To validate this concept, we examined the relationship between cost-per-success and minimum/maximum behavior change probabilities (i.e., $p_{\min}$ and $p_{\max}$) across various incentive mechanisms.

As illustrated in Figure 10, the fixed incentive consistently incurred a cost of 50 KRW per success, regardless of behavior change probabilities. In contrast, the random and personalized incentives demonstrated greater cost-efficiency as behavior change probabilities increased. To explore this further, we conducted a regression analysis with cost-per-success as the dependent variable and minimum/maximum behavior change probabilities as independent variables.

The analysis revealed that increasing the minimum behavior change probability significantly reduced cost-per-success in the personalized incentive mechanism ($\beta = -66.00$) compared to the random incentive mechanism ($\beta = -30.58$). Conversely, raising

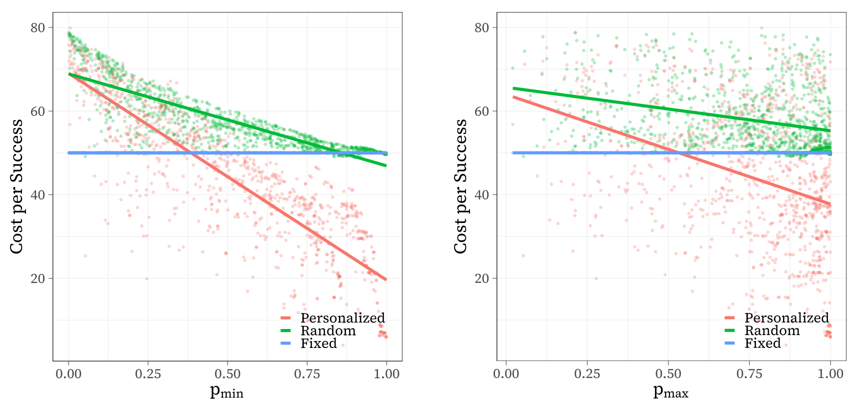
Figure 9: Comparisons of different incentive mechanisms in the simulation study



the maximum behavior change probability increased the cost-per-success more in the personalized incentive mechanism ($\beta = 31.80$) than in the random incentive mechanism ($\beta = 16.50$).

These findings suggest that personalized incentive mechanisms can be particularly cost-effective when users are likely to change their behavior with a smaller incentive. Additionally, the personalized approach can still motivate behavior change among users who strongly prefer higher incentives by offering a suitably larger amount. Thus, we can conclude that our personalization mechanism operates as intended, effectively balancing cost expenditure with the encouragement of behavioral changes.

Figure 10: Relationships between cost-effectiveness and behavior change probabilities in the simulation study



(a) The minimum probability of behavior change, $p_{\min}$

(b) The maximum probability of behavior change, $p_{\max}$