# GoldenTime: Exploring System-Driven Timeboxing and Micro-Financial Incentives for Self-Regulated Phone Use

Joonyoung Park
Graduate School of Knowledge Service Engineering, KAIST
Daejeon, Republic of Korea
zelatore85@kaist.ac.kr

Hyunsoo Lee
Graduate School of Knowledge Service Engineering, KAIST
Daejeon, Republic of Korea
hslee90@kaist.ac.kr

Sangkeun Park*
Graduate School of Knowledge Service Engineering, KAIST
Daejeon, Republic of Korea
sk.park@kaist.ac.kr

Kyong-Mee Chung
Yonsei University
Seoul, Republic of Korea
kmchung@yonsei.ac.kr

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

## ABSTRACT

User-driven intervention tools such as self-tracking help users to self-regulate problematic smartphone usage. These tools basically assume active user engagement, but prior studies warned a lack of user engagement over time. This paper proposes *GoldenTime*, a mobile app that promotes self-regulated usage behavior via system-driven proactive timeboxing and micro-financial incentives framed as *gain* or *loss* for behavioral reinforcement. We conducted a large-scale user study (n = 210) to explore how our proactive timeboxing and micro-financial incentives influence users' smartphone usage behaviors. Our findings show that *GoldenTime*'s timeboxing based micro-financial incentives are effective in self-regulating smartphone usage, and incentive framing has a significant impact on user behavior. We provide practical design guidelines for persuasive technology design related to promoting digital wellbeing.

## CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

## KEYWORDS

Digital Wellbeing, Smartphone Intervention, Self-regulation, Financial incentive, Timeboxing

**ACM Reference Format:**
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 *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3411764.3445489

*Corresponding authors: Sangkeun Park (Evaluation) and Uichin Lee (Design and Discussion).


*CHI '21, May 8–13, 2021, Yokohama, Japan*

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445489

## 1 INTRODUCTION

Behavior changes typically start with small, meaningful changes as popular habit formation books recommend [18]. As one of the serious societal issues, a lack of self-regulated usage of smartphones often involves frequent excessive usage or checking [35, 38], which considerably distract users' everyday lives [31]. It would be important to break existing problematic usage habits while simultaneously establishing healthy ones. Goal setting theories for behavior changes showed that setting specific and proximal goals with progress feedback and rewards would greatly increase the likelihood of goal achievement [47, 65]. The same principle can be applied to smartphone usage regulation.

Problematic usage behaviors are associated with a lack of time management skills [20, 33, 68]. Well-known time management techniques would be using timeboxing which allocates a fixed time slot (called a timebox) during which planned activity will be performed. Smartphones as distractors in most cases can be considered in timeboxing rules such as limiting their usage for a given time slot. Self-limiting features have been widely supported in both commercial apps and research prototypes [50].

Existing self-limiting methods mostly presume user-initiated timeboxing whenever a user must enable self-limiting features to block potential distraction. While this approach offers a great autonomy to the user, prior studies warned a lack of user engagement over time [32, 49]. Designing sustained intervention engagement is critical for successful behavioral change. Recent studies explored several proactive (or system-initiated) intervention strategies with event-based triggering such as unlock event [34], app usage [67], and location [31]. As goal setting theories recommend, prior studies also leveraged behavioral reinforcement strategies [16] such as token bestowal (e.g., badges), social recognition (e.g., recognizing progress), and tangible rewarding (e.g., prize or money). In particular, financial incentives are generally considered as effective behavioral reinforcement [23].

In this work, we design and evaluate *GoldenTime* to explore the design space of system-driven timeboxing with micro-financial incentives for sustainable intervention engagement. A combination of timeboxing and micro-financial incentives creates a unique opportunity for reinforcing "small changes" (e.g., for a given hour, use your phone less than 10 minutes). Furthermore, we experiment with "incentive framing" (e.g., gain or loss) to understand how framing affects continuous decision making for behavior changes. We iteratively prototyped *GoldenTime* and conducted a randomized controlled field trial (control, gain, and loss groups) with n =210 for four weeks. We conducted a log-data analysis to show the effectiveness (e.g., smartphone usage time, and successful timeboxes). In addition, we investigated user experiences of system-driven timeboxing with and without micro-financial incentives. Our results showed that micro-financial incentives, when framed as loss, were more effective than the other cases. Furthermore, the overall process of behavior evaluation and coping differs significantly, ranging from the demotivation process to the devaluation and revaluation process. Our findings helped us to explore several practical design directions, such as considering context-awareness, exploring micro-financial incentive design space, and enabling data-driven actionable insights for self-reflection. The key contributions of our study are as follow:

- We designed and implemented *GoldenTime*, a smartphone intervention app for promoting self-regulated usage behavior. *GoldenTime* was designed based on behavioral economics, self-regulation theory, and the results of prior empirical studies on smartphone regulation.
- We quantitatively and qualitatively analyzed how different framing of a micro-incentive mechanism has a different effect on self-regulation over time by performing a large-scale between-subjects experiment in the wild (n = 210).
- We provided empirical findings and design implications on how smartphone interventions based on timeboxing and micro-financial incentives could help users to self-regulate their phone usage.

# 2 THEORETICAL BACKGROUNDS AND RELATED WORK

In this section, we include an overview of self-regulation theory and review how the theory has been applied to the prior studies on intervention design for smartphone usage regulation. In addition, we review how financial incentives and framing effects have been utilized in prior studies on behavior change.

## 2.1 Smartphone Usage Regulation with Self-Regulation Theory

Recent HCI studies have introduced various intervention systems that aim to regulate problematic smartphone usage (e.g., [24, 34, 37]). One of the most commonly used mechanisms to regulate smartphone usage is inducing self-regulated usage behavior by increasing self-awareness of usage behavior. The mechanism is grounded in self-regulation theory (SRT) [4], and existing studies have tried to verify the effectiveness by researching various intervention designs based on smartphone use as a kind of self-regulation process. Self-regulation is a process of improving one's behavior through unconscious or conscious efforts to achieve a target behavior to obtain a result that matches one's original resolution. Bandura's social cognitive theory [4] argues that self-regulation can be achieved through the following three stages—1) self-observation, 2) self-judgment, and 3) self-reaction. *Self-observation* entails observing or tracking one's own behavior. *Self-judgment* then evaluates the observed behavior based on individual, social and collective norms. If the behavior does not fulfill or meet desired standards, the behavior is corrected through the self-reaction process, leading to an action that matches the desired goal.

Prior smartphone intervention studies have pointed a lack of self-regulation as the main cause of one's problematic behavior [40, 41] Based on this, they proposed intervention designs and tools to induce regulated usage by improving self-regulation. One of the key intervention methods is "usage tracking and visualization." [48, 58, 64] This feature facilitates self-monitoring by statistically visualizing a user's phone usage history (e.g., daily usage time, app usage frequency, etc.). Usage tracking and visualization help users to recognize their own usage behavior. When users exceed their predefined usage target limits, they provide feedback such as warning messages [24] or vibration [57] to enhance awareness. Monitoring and visualization techniques were used to promote mindfulness of use behavior through self-monitoring support.

Recently, HCI researchers have begun to study intervention designs that induce self-regulated behavior in smartphone use in a more restrictive and coercive way. The studies explored the effects of the "uncomfortable interactions" [5], a design mechanism that suggests that deliberately introducing discomfort to interactive experiences can be an important design tool that induces positive long-term goals. Studies that leveraged this concept have shown that minor discomfort or micro-boundary introductions [13] when users launch their phones can induce self-regulated behavior by increasing the user's mindfulness. A variant of this study is temporary blocking or restriction of smartphone usage based on a user's predefined rules (e.g., time/physical activity [46], location-based blocking [31]). For example, Kim et al. [32] studied the effects of self-interruption management tools that restrict access to predefined non-productive applications in multi-device environments. The study designed "PomodoLock," a software tool that utilizes timeboxing to manage self-interruption in concentration mode initiated by the user. They verified that PomodoLock improves users' time management skills by reducing their self-interruption. Another intervention mechanism proposed by Kim et al. [34] intentionally redirects usage behavior by inserting a lockout task when an app that has been previously classified as a black app is launched. This lockout task is an intentional inserted "gulf-of-execution" to delay the phone usage interaction. From the perspective of dual process theory, this design can be seen as a mechanism to induce self-reflection and judgment by moving the decision-making of usage behavior from System 1 to System 2, and effects of self-regulation have been proven through existing empirical studies.

There were also studies on smartphone usage intervention using social tools and social support concepts [36–38]. Ko et al. used a strategy to induce regulated behavior by setting and sharing targets for limiting use among group members [38]. The study observed that users could come to an objective realization of their usage behavior and maintain the target behavior through the process of socially comparing and competing with their usage behavior (e.g., usage time). Lock n'LoL [37] is an intervention app designed to explicitly share usage time between group members, limiting the smartphone usage behavior together, and focusing on group activities. The app induces self-regulated behavior by providing social awareness about usage behavior through this "synchronous restriction." There were also studies that used social support in parent-child interaction [25, 36]. FamiLync [36] supports regulated behavior by using a public space where parents and children share usage information. Plan & Play enables self-regulation through a function that allows parents and children to plan and set goals for using the app together, and visualize their goal-setting [25]. The study showed that social support based on social learning and competition has a positive effect on smartphone usage regulation.

Prior smartphone intervention studies have revealed that the user's engagement with the intervention is crucial to sustaining the self-regulated usage behavior. According to these prior studies, one caveat of restrictive intervention is weakened user engagement caused by user fatigue [34], as it inflicts coercion upon users. In addition, passive monitoring and visualization tools have also been mentioned as factors that weaken the intervention effects [10]. Through empirical research, they emphasized the necessity of an intervention design that helps users to observe and recognize their usage behaviors and to maintain regulated behaviors [24, 37, 38]. In addition, existing persuasive interaction design studies mentioned that behavioral change is not a single action, but a process, emphasizing that it is a process of achieving the ultimate target behavior through micro-behavior changes [17, 56].

Based on the aforementioned background and the need for further research, we propose *GoldenTime*, a mobile app that promotes self-regulated usage behavior via system-driven proactive timeboxing and micro-financial incentives framed as *gain* or *loss* for behavioral reinforcement. *GoldenTime* leverages a timeboxing mechanism to enable real-time tracking of user's self observation and recognition of their smartphone usage behavior (i.e., micro-usage behavior). In addition, the app provides financial incentives to reinforce regulated behavior, thereby making users realize the monetary value of their usage behavior. Here, we define a mechanism that provides financial incentives for the user's self-regulated usage behavior (i.e., micro-usage behavior) in the time boxing-based proactive intervention process as a "micro incentive mechanism."

## 2.2 Financial Incentives for Behavior Change

*2.2.1 Gain-Framed Incentives vs. Loss-Framed Incentives.* In an attempt to enhance self-regulation and reinforce positive behavior change, adding financial incentives to reward/punishment features in behavior change tactics or systems have been widely discussed in various contexts [50]. Financial incentives generally grant tangible rewards to decrease/abstain certain behaviors and increase/induce desired behavior [23]. These contingent incentives can be either framed as *gain* or *loss*.

According to behaviorist theory [66], *gain-framed* incentives are related to positive reinforcement, offering a reward once a desired behavior/outcome is accomplished. This incentive scheme has been shown to be effective in diverse realms of behaviors (e.g., smoking cessation [69], physical activity [52], and weight loss [22]). In HCI studies, awarding virtual badges as an incentive to reinforce behavior maintenance [19] or increase participation in social computing contexts [3] has been deployed. Agapie et al. [1] also experimented on awarding "cheat points," in which badges are given for sustained use of a system. Other commonly studied contexts include crowd-sourced works and management [53] and proximal health interventions [2, 45], where offering an incentive has shown to be effective in inducing desired behavior.

Although most incentive schemes tend to offer gain-framed incentives, inducing some feeling of *loss* in case of failure may be more motivating than rewarding the same amount (i.e., loss aversion) [28]. This *loss-framed* incentive, a negative punishment that imposes a sense of loss or penalty once the desired outcome is not accomplished, has shown that framing the outcome of incentives as a loss can increase initiation or change of behavior than gain-framed incentives in inducing one's positive behavior change (e.g., smoking cessation [63]). While there have been wide applications and attempts that leverage gain-framed incentive in diverse behavior realms, HCI studies that exploit loss-framed incentives have been relatively unexplored. Rather, studies have suggested some concerns that a sense of loss may invoke a user's attrition and abandonment of a behavior intervention system [9, 11].

Overall, to the extent of our knowledge, there was a lack of studies that quantitatively and qualitatively investigated how users perceive different incentive frames and how these design choices affect the way users interact with a behavior change intervention system. With *GoldenTime*, we seek to evaluate a thorough user experience through the lens of the self-regulation process and how these different frames affect one's problematic smartphone usage.

*2.2.2 Micro-Financial Incentives for Behavior Change.* Despite the well-known benefits of the aforementioned incentive schemes, these incentives are typically coarse-grained and distributed in bulk payments under a single milestone (e.g., at the end of research, upon the accomplishment of a target behavior) [55], making it difficult to track real-time changes in one's behavior or make flexible adjustments along the way. To combat such challenges, offering flexibility in terms of payment time, incentive types and the amount have emerged. This approach, also known as "Micro-Financial Incentives," sets several milestones for a target behavior (e.g., productivity, physical activities) to encourage people to reach the next milestone and reward people.

Recently, this flexible approach is being increasingly deployed in mobile-phone based user studies (e.g., awarding $1 for answering survey questions [12], participatory sensing [62]). Yamabe et al.'s study [70] also presented the effectiveness of activity-based micro-financial incentive mechanisms to discreetly steer user behaviors toward desired patterns. Such results from past studies highlight a room for future research, suggesting micro-financial incentives as an important design feature that helps to create sustainable habit formation [62].

Despite such reported benefits, explorations on design spaces that leverage micro-financial incentives for positive behavior change have largely been unexplored. Particularly, to the best of our knowledge, our study is the first HCI research that attempts to understand the effectiveness of micro-financial incentives in conjunction with the usage of a system-driven timeboxing strategy. With our approach, we aim to study how a behavior intervention system that leverages different incentive framing affects smartphone usage and how micro-incentive based design influences overall user experience in terms of positive behavior reinforcement. Towards these goals, we set the following research questions:

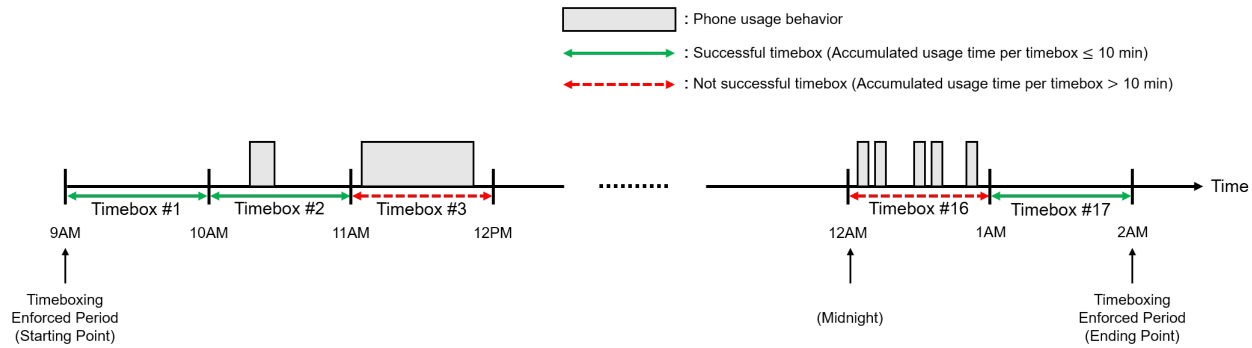



**Figure 1: System-Driven Timeboxing Operation**

- **RQ1:** How do we design system-driven (or proactive) intervention that promotes micro-behavior changes with micro-financial incentives?
- **RQ2:** Does micro-financial incentive based proactive intervention and its framing have positive effects on self-regulating smartphone usage?
- **RQ3:** How do system-driven intervention and incentive framing influence usage behaviors and self-regulation strategies over time?

## 3 GOLDENTIME SYSTEM DESIGN (RQ1)

In this section, we elaborate on the design components and the overall design process of *GoldenTime* to answer our following research question: *RQ1: "How do we design system-driven (or proactive) intervention that promotes micro-behavior changes with micro-financial incentives?"*

### 3.1 Exploring Design Space

We aimed to explore design space of proactive intervention on smartphone overuse with financial rewards. We went through an extensive literature review of diverse realms of studies that include HCI, psychology, and behavioral economics. As a result, we came up with three main design components: *timeboxing*, *micro-incentive mechanisms*, and *intervention operations*.

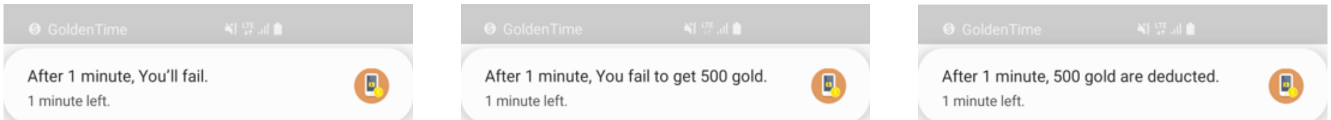
*3.1.1 Timeboxing Design.* We considered an intervention method that allows users to manage their smartphone usage time and behavior by applying *timeboxing*, which is a time management technique that manages tasks or plans using a fixed timebox. As to timeboxing design, we considered its period, time limit, as well as initiative, and operating time to enable users' continuous self-observation and recognition of their smartphone usage behavior.

**System-Driven Timeboxing**. As mentioned above, it is critical to design an intervention system that sustains user engagement for effective control behavior. Despite the reported effects of the timeboxing technique, prior studies have also suggested that user-initiated timeboxing (e.g., manual operation and scheduling [32]) has a limited impact on the maintenance of the desired behavior, as users can arbitrarily stop their behavior anytime. To provide effective intervention, we adopted a system-driven timeboxing, which automatically starts timeboxing regardless of users' intentions. System-driven timeboxing is designed to be initiated and updated every hour (e.g., 9 AM, 10 AM) to track users' smartphone usage time and behavior. We initially set our timeboxing activation hours from 9AM to 6PM. Our first design considered the working hours to minimize work distraction caused by smartphones. As shown later, a pilot study was later conducted to find the appropriate activation hours. We decided to uniformly apply this timeboxing period to cover major daily activities; most phone usage occurs during working time and before sleep [14], and thus, applying this uniform period could capture major usage patterns. The finally designed system-driven timeboxing is shown in Figure 1.

**Timebox Duration**. We considered the timebox duration (period) to support effective regulated usage. We set the timebox duration to one hour. This design was from prior studies that empirically used similar duration [31, 37, 38] as highlighted in Ko et al.'s work [38] on the time unit. Through the online survey, this study emphasized the need for a design that limits the usage for a certain period of time according to the users' activity time, since limiting the usage time per day without the consideration of a user's context may disturb a user's primary task; e.g., playing games in a classroom. Furthermore, a social support-based smartphone intervention study [37] showed that users tended to plan their activities or tasks on an hourly basis and preferred automatic recharging of usage time allowance on an hourly basis to offer flexibility in limiting usage behavior.

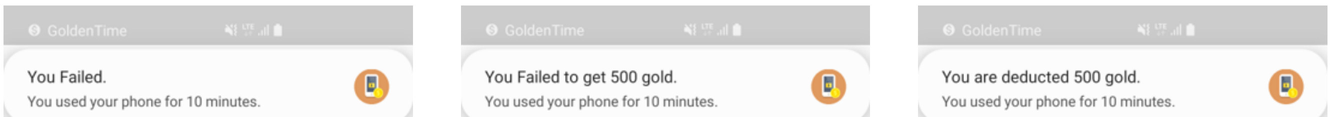
**Usage Time Limit**. We designed the time limit in timeboxing to prevent too much time being distracted from by the primary tasks. We set the time limit to 10 minutes. This setting was from prior studies that used time limits 5 minutes [31, 37] or 10 minutes [38] per hour. Prior studies have empirically shown that such time restriction is effective in promoting regulated behavior and increases the flexibility of the intervention design as well. However, since each study had different target behavior and context (e.g., studying in the classroom [31] or social activities [37, 38]), we decided to evaluate and finalize our configuration through a pilot study.

*3.1.2 Micro Incentive Mechanism.* For incentives design, we considered the amount of an incentive and its framing as follows:




**(a) Control Group** **(b) Gain Group** **(c) Loss Group**

**Figure 2: Warning Messages**




**(a) Control Group** **(b) Gain Group** **(c) Loss Group**

**Figure 3: Fail Messages**

**Financial Incentive Reward**. We used a virtual concept of a coin, called *Gold*, as an incentive unit for our micro-incentive mechanism. *GoldenTime* rewarded 500 Golds when the user followed the time limit rule (i.e., 10 minutes) in a given time box. 500 Golds have a monetary value of approximately 5 cents in USD. This amount was chosen based on our pilot study.

**Incentive Payment Framing**. We referred to existing studies on financial incentives for behavior changes (e.g., weight loss, smoking cessation [26, 51]) that leveraged the framing effect of incentives (gain vs. loss) [52, 63]. Framing was set as follows. In the gain framing, users were rewarded with 500 Golds for each successful time box. In the loss framing, users were given a fixed total amount of gold at the start day. If a user failed to regulate his/her smartphone usage time, 500 Golds were deducted for each failed time box. Total payment was identical for both the gain and loss groups. The maximum incentive amount for the three-week intervention period was approximately 16 USD.

## 3.2 Intervention System Design

Combining system-driven timeboxing and micro-financial incentive design, we designed the *GoldenTime* intervention system. In this process, two different micro-incentive mechanisms were designed. We hypothesized that the loss framing would be more effective in reinforcing the self-regulated smartphone usage behavior. This hypothesis was derived from the *endowment effect* and *loss aversion* of behavioral economics [28].

*3.2.1 Intervention Operation.* We considered a smartphone intervention mechanism that operates under the system-driven timeboxing. Existing empirical studies have emphasized the importance of continuous self-monitoring to maintain regulated usage behaviors and notifications that can remind users of their target behaviors [31, 67]. According to the behavioral intervention technology model (BIT) [54], this is a behavioral strategy related to monitoring and feedback, which enables users to understand and realize their current state (i.e., the behavior of using smartphones) to achieve the target behavior.

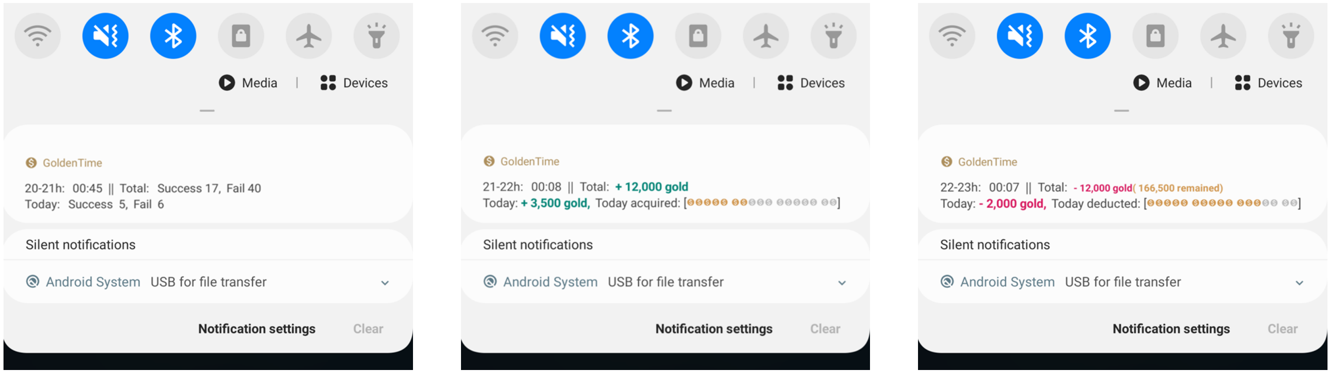
*3.2.2 Intervention Feedback.* We designed an intervention mechanism that provides notification based on the 10-min time limit in the system-driven timeboxing design. Specifically, a notification message was used to warn a user of the usage behavior, when the accumulated usage time elapses 9 minutes (notifying "1 min left") and when the usage time elapsed 10 minutes (failure notification). In these two notifications, vibration was sent together to maximize the effect of notification. A prior study found that vibration can be a "nudge" in the intervention process [57].

We designed a specific warning message phrase to provide notification messages, which were designed differently according to an incentive frame. The purpose of a notification message was to provide warning of regulated usage behavior failure through micro-financial incentives. Therefore, we designed a warning message as failure of micro-incentive acquisition (gain framing) or deduction of micro-incentive (loss framing). In other words, the gain mechanism provided a warning message for gold acquisition failure due to regulation failure, and the loss mechanism provided a prepaid gold deduction warning message to the user. The notification messages with two framings are shown in Figure 2 and 3.

## 3.3 Iterative Design Process

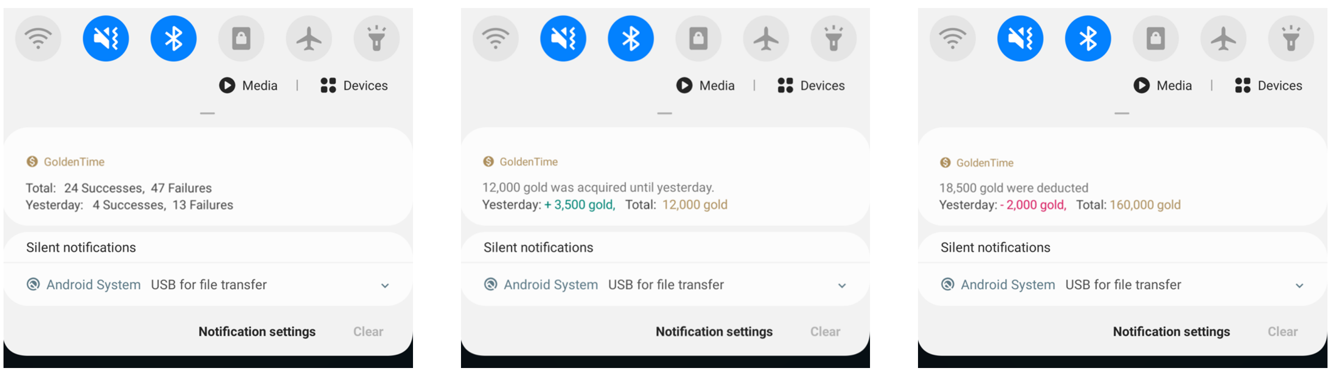
We performed rapid iterative prototyping to finalize and implement the *GoldenTime* mobile app design. All the pilot studies were performed in-the-wild (1st: 3 days, 2nd: 4 days, 5 days for each round, 3rd: 2 weeks). In the initial pilot study, four participants (two females, age: M = 25.50, SD = 1.66) were asked to test a paper prototype and provide feedback on overall interface design.

Two pilot studies with the high-fidelity prototypes were conducted and surveys and interviews were used to assess timeboxing features and appropriateness of an incentive amount. A total of 12 users (one female, age: M = 24.50, SD = 2.90) participated in this experiment. In the exit survey, most of the participants (n = 10, 83%) evaluated that the system-driven timeboxing design was effective in regulating usage behavior (M = 3.78, SD = 0.85). They mentioned that warning notifications under the system-driven timeboxing are convenient as the design feature allows automatic and continuous tracking of their usage behavior. As to timeboxing operating




**(a) Control Group** **(b) Gain Group** **(c) Loss Group**

**Figure 4: Real-time Notifications**




**(a) Control Group** **(b) Gain Group** **(c) Loss Group**

**Figure 5: Daily Notifications**

hours, eight participants mentioned that it needed to be adjusted, considering that most problematic phone use was concentrated from evening to dawn (before sleep). From the feedback, we determined to adjust the timeboxing hours from 9AM to 2AM (next day). Meanwhile, 11 participants evaluated that the micro-incentive amount level and unit setting were appropriate for reinforcing their self-regulated usage behavior (M = 4.17, SD = 0.60).

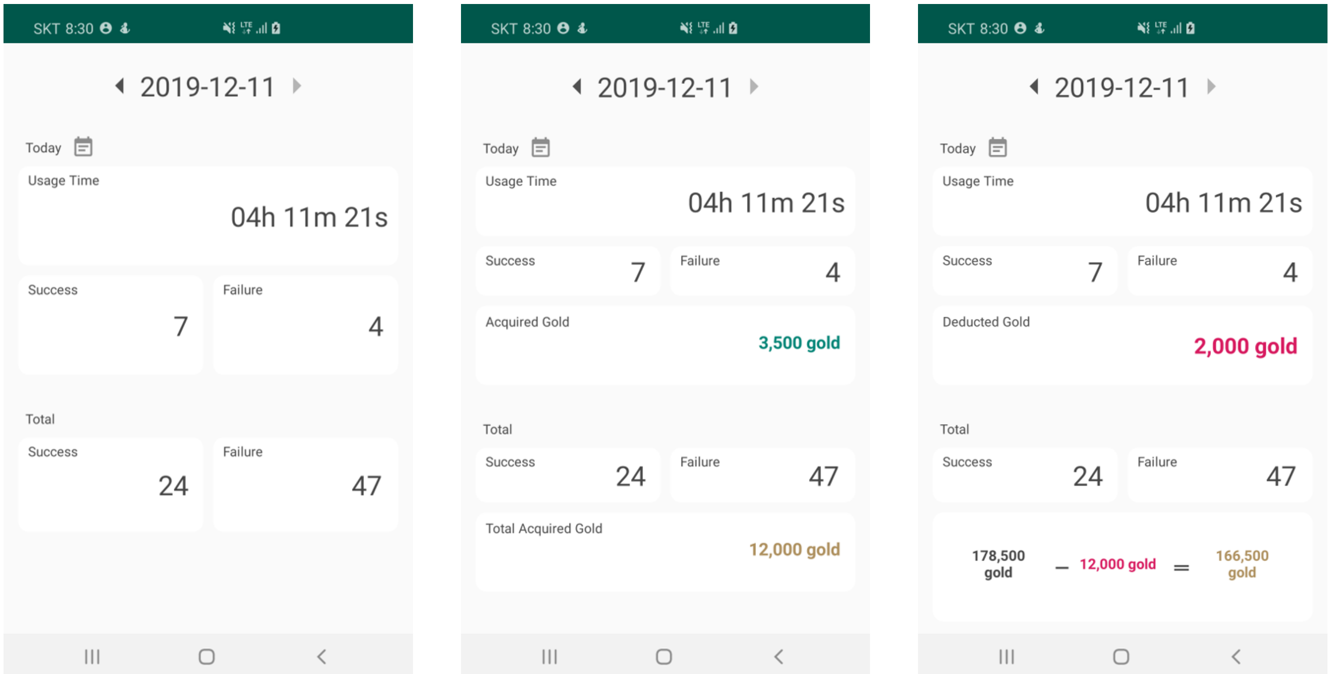
The final pilot study was conducted with six participants (three females, age: M = 24.83, SD = 1.77) to evaluate the stability of the final prototype. As a result, the following design change was made after three pilot studies: intervention periods (9 AM – 6 PM to 9 AM - 2 AM).

## 3.4 GoldenTime System Description

We implemented the *GoldenTime* system based on our design exploration. *GoldenTime* is composed of 1) Warning Notification Feedback, 2) Real-time Usage Feedback, 3) Daily Notification Feedback, and 4) Usage Statistics Dashboard in terms of system components. For a field experiment, we built three different *versions* of the *GoldenTime* system: 1) control, 2) gain framing, and 3) loss framing. *GoldenTime* runs on Android phones with operating system version 8.0.0 (Oreo) or higher.

*3.4.1 Warning & Fail Notification Feedback.* As shown in Figure 2, *GoldenTime* provides an hourly-based pop-up message that warns users of the elapsed usage time, 9 and 10 minutes, respectively. The pop-up message included different content for each of the three app versions. Figure 3 shows that the prototypes for the two experimental groups (i.e., gain and loss) included incentive acquisition failure or deduction statements, and the control prototype included regulation failure warning. All versions of warning and fail notification feedback were accompanied by vibration feedback as well.

*3.4.2 Real-time Notification Feedback. GoldenTime* supports users to check their usage statistics real-time via Real-time Notification Feedback (see Figure 4). Real-time Notification Feedback displays a user's usage time (i.e., elapsed time) within the timebox using the Android notification bar. To display information related to regulated usage behaviors, daily and overall statistics of timeboxing were presented. Timeboxing status was updated every hour. Real-time Notification Feedback for the control group was given information




**(a) Control Group** **(b) Gain Group** **(c) Loss Group**

**Figure 6: Dashboard**

on the current status of successes and failures (i.e., counts of each instance). In addition, the Real-time Notification Feedback for the incentive groups was shown the amount of incentives earned or deducted up to date.

*3.4.3 Daily Notification Feedback.* Users received feedback on the amount of usage time a day before at the beginning of daily timeboxing (i.e., 9AM) (see Figure 5). In this process, the three versions provided different Daily Notification Feedback. The Daily Notification Feedback of the control group provided information on the number of regulation success and failure, and the Daily Notification Feedback of the two incentive groups provided information on the status of incentives. Gain framing's Daily Notification Feedback showed the total amount of incentives earned up to yesterday and the amount of incentives earned during the day yesterday. Loss framing's Daily Notification Feedback showed the total amount of incentives deducted up to yesterday and the amount of incentives deducted during the day yesterday.

*3.4.4 Usage Statistics Dashboard. GoldenTime* provided Usage Statistics Dashboard, a visualization interface for users to check usage statistics (see Figure 6). The usage statistics information basically included the daily usage time and the number of regulation success and failure. The version of the Usage Statistics Dashboard for both incentive groups also included incentive (i.e., gold) status information. The version of the Usage Statistics Dashboard for Gain framing showed the amount of incentive earned, and the Usage Statistics Dashboard for Loss framing showed the incentive amount deducted from the total amount of incentives paid in advance. *GoldenTime* supports users to navigate past usage records by using the Android's Calendar control.

## 4 EXPERIMENT DESIGN

We conducted a 4-week in-the-wild deployment experiment with 210 participants to investigate how the proposed *GoldenTime* app affects self-regulated usage behavior. Specifically, our main concerns are 1) whether micro-incentive based proactive intervention and its framing have effect on self-regulating smartphone usage (RQ2), and 2) how system-driven intervention and incentive framing influence usage behaviors and self-regulation strategies over time (RQ3).

### 4.1 Methods

We tried to figure out how different micro-incentive mechanisms have different effects on regulated usage behavior. To this end, we conducted a randomized controlled trial. We designed a between-group experiment by randomly assigning the participants into control group and two experimental groups (i.e., gain and loss). We provided the control with a version of the *GoldenTime* app to which the micro-incentive mechanism was not applied. Since this app did not provide micro-incentives, we did not include any incentives information in all UIs. Instead, the notification message contained some warning phrases to alert regulation failure.

The experiment was conducted for 4 weeks with 1 week baseline period and 3 weeks of the intervention period. All three groups received no intervention during the baseline period, and they experienced different intervention methods during the intervention period. To compare smartphone usage behavior before and after the intervention, we collected smartphone usage log data for all the periods. After the experiment, we performed a follow-up survey and semi-structured interview to understand their experience using the *GoldenTime* app. The interview questionnaire was composed of questions about the effects of *GoldenTime*'s features (e.g., micro-incentives mechanism) on self-regulation and its reason. The purpose of the exit survey was basically to evaluate the usability of *GoldenTime*. We also included the interview questionnaire items into the survey items to analyze the intervention experiences of those who did not participate in the interview. Our study was approved by the university's Institutional Review Board (IRB) and was conducted with participants' written consent.

## 4.2 Participants

We first calculated the sample size to ensure the statistical significance of the experimental results. To compare and analyze the mean difference for each period of the three groups, we set the effect size to 0.25 (i.e., Cohen's f medium effect size), and the significance level and power value were set to 0.05 and 0.95, respectively. The sample size calculated using G*Power [15], the most reliable tool used to calculate statistical power was 159. We decided to recruit more than 200 participants, factoring a dropout rate of 10–20%. The scale is similar to the existing financial incentive experiment studies [59].

We recruited participants through online communities and portals of two large universities. To determine the precise selection criteria, we used the transtheoretical model (TTM) [61] to describe the phase of behavior change. We chose those volunteers whose TTM levels were stage two (considering reduction) or stage three (ready to reduce smartphone use) in the recruitment. After this initial screening, we randomly assigned selected participants.

From a total of 284 volunteers, 223 of them were screened through the screening process (77 females, age: M = 24.07, SD = 2.88) and selected as participants. Our demography analysis revealed that no statistical significance was observed across three different groups in terms of 1) mean age: F(2, 222) = .336, $p$ = .715, and 2) sex ratio: $\chi^2(2, N = 223) = .18, p = .991$.

## 4.3 Procedure

We had an orientation to explain how to proceed with the experiment. Participants were briefed on the experimental guidelines, including a description of the data to be collected during the experiment. We had the following cases of dropouts in the orientation: 1) iPhone users (n = 7) were excluded as they falsely reported their device as Android for the participation, and 2) six refused to join due to privacy concerns (e.g., usage tracking). After the orientation, 210 participants finally joined the experiment.

We collected participants' smartphone usage logs during all periods. To confirm the data collection rate, we checked the database in real time to check the participants' data collection progress. In this process, we tried to resolve the problems by sending emails and messages to contact individual participants who had issues with data collection. After the end of the experiment, all participants were asked to perform an exit survey for the *GoldenTime* app experience. In addition, we randomly selected 30 participants (10 per group) and conducted an in-depth interview on their experiences.

During the 4-week experiment, a total of 141,120 timebox instances were collected from 210 participants. In the process of data analysis, we found 5 participants whose usage logs were very unusual (e.g., no phone usage for a week). We regarded them as outliers and excluded them from data analysis. As a result, we included only the data of a total of 205 participants (Control: n = 68, Gain: n = 68, Loss: n = 69).

## 4.4 Data Analysis Methods

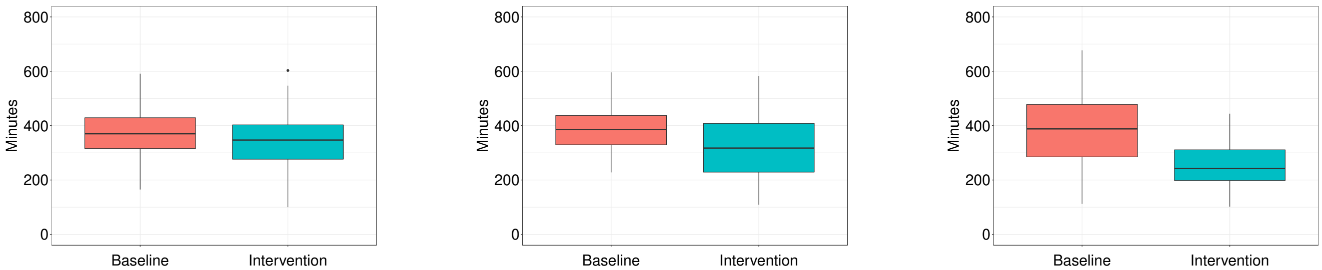
The purpose of our study was to compare and analyze how different micro-incentive mechanisms have different effects on self-regulation over time. In order to quantitatively compare the three mechanisms (along with the baseline), we used two metrics from the collected smartphone logs: a daily average of 1) smartphone usage time and 2) regulation success count.

In our experiment, the intervention was given only during timeboxing operating hours (9AM–2AM). When calculating the regulation success count, we decided to consider the entire day (i.e., beyond the timeboxing operating hours), because we wanted to understand how *GoldenTime* affected phone usage throughout the day. To measure the number of regulation success, we counted cases where the usage time was less than 10 minutes per hourly timebox. Since our timeboxing intervention was based on the time limit of using less than 10 minutes per hourly timebox, we regarded such a case as regulation success. For a given day, the regulation success count could be at most 24. For statistical analyses, we used a mixed ANOVA to determine whether any change in daily smartphone usage behavior measured by the metrics is the result of the interaction between the group (i.e., Control, Gain and Loss; that is, the intervention type, which is the between-subjects factor) and period (i.e., the within-subjects factor, consisting of four-time periods).

We also tried to understand how user experiences for different micro-incentive mechanisms had effects on the regulation. Through our thematic analysis of interview data and exit survey [7], we investigated the following: 1) overall experience using *GoldenTime*, 2) reflection on problematic smartphone use, and 3) changes after using *GoldenTime*. Two authors collaboratively performed content analysis using ATLAS.ti Cloud. In this process, each author performed creating codes for each point. Thematic analysis using affinity diagramming was performed repeatedly until all authors reached consensus.

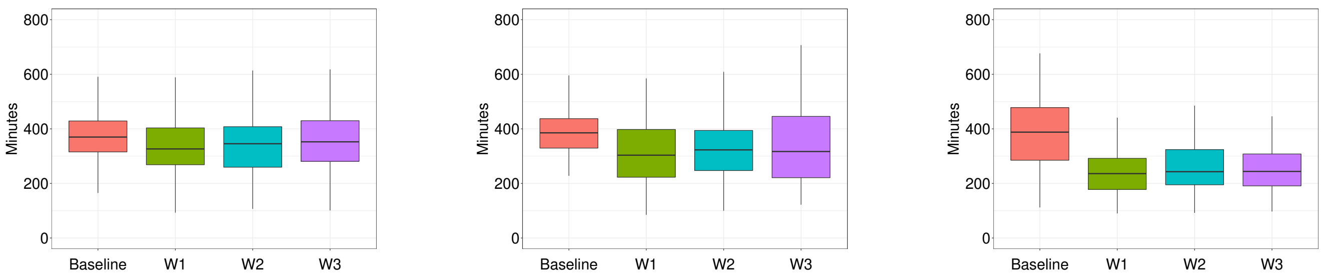
# 5 RESULTS

We describe the experimental data analysis results to answer our research questions: *RQ2: Does micro-financial incentive based proactive intervention and its framing have positive effects on self-regulating smartphone usage?* and *RQ3: How do system-driven intervention and incentive framing influence usage behaviors and self-regulation strategies over time?* Exit survey and interview data were analyzed along with smartphone usage log.




(a) Control Group (b) Gain Group (c) Loss Group

**Figure 7: Daily Average of Smartphone Usage Time (Baseline vs. Intervention)**




(a) Control Group (b) Gain Group (c) Loss Group

**Figure 8: Daily Average of Smartphone Usage Time by Period**

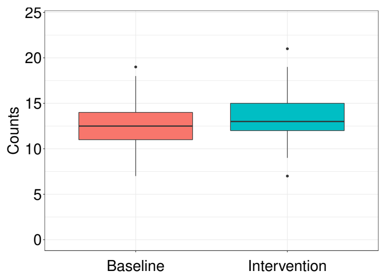
## 5.1 Effects on Regulated Usage (RQ2)

As mentioned in section 4.4, we compared the differences between groups over time using two metrics: 1) a daily average of smartphone usage time (SUT), and 2) a daily average of regulation success count (RSC).

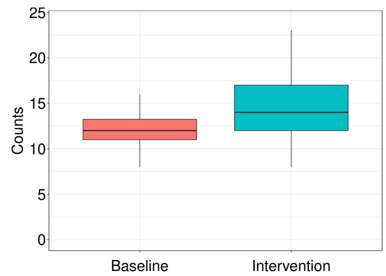
*5.1.1 Daily Average of Smartphone Usage Time (SUT).* We first compared and observed how SUT was changed for each group to investigate the intervention effect. Descriptive statistics showed that SUT overall decreased during the intervention period in all three groups as shown in Figure 7 and 8; Control (Baseline: M = 374.56, SD =101.53; Intervention W1: M = 333.07, SD = 101.03; Intervention W2: M = 339.69, SD = 100.08; Intervention W3: M = 356.43, SD =116.32), Gain (Baseline: M = 390.97, SD =82.85; Intervention W1: M = 310.06, SD =129.84; Intervention W2: M = 321.37, SD = 120.83; Intervention W3: M = 326.75, SD = 135.30), and Loss (Baseline: M = 387.80, SD = 141.96; Intervention W1: M = 242.65, SD = 89.30; Intervention W2: M = 257.09, SD = 92.36; Intervention W3: M = 258.06, SD = 88.43).

A mixed ANOVA with a Greenhouse-Geisser correction determined that there were significant differences across the four-time periods, F(2.27, 458.21) =77.74, $p$ < .001, $\eta^2$ = .28, and significant differences between groups, F(2, 202) = 8.96, $p$ = .001, $\eta^2$ = .08, in SUT. There was also a significant interaction effect between period and group, F(4.54, 458.21) = 11.05, $p$ < .001, $\eta^2$ = .10. Following up this interaction indicated that SUT of control was significantly reduced from baseline to the first treatment week (Mean difference = 41.49, SE = 12.79, $p$ = .008) and the second treatment week (Mean difference = 34.87, SE = 12.75, $p$ = .041). However, the mean difference with the third treatment week (Mean difference = 18.13, SE = 14.02) was not significantly different from the baseline. SUT of two incentive groups decreased at all treatment periods compared with the baseline, all of which were significantly different. In the Gain, mean difference compared with the baseline was 80.91 (SE = 12.79, $p$ < .001) at the first week of treatment, 69.60 (SE = 12.75, $p$ < .001) at the second week of treatment, and 64.22 (SE = 14.02, $p$ < .001) at the last week of treatment. In the Loss, each mean difference was 145.15 (SE =12.69, $p$ < .001), 130.71 (SE = 12.65, $p$ < .001) and 129.74 (SE = 13.92, $p$ < .001), respectively.

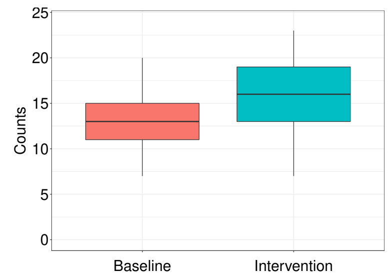
We further performed post hoc analysis of multiple comparisons of each group by period using the Bonferroni correction. The analysis revealed that there was no significant mean difference between groups at baseline. However, there were significant differences between the Loss and others in all treatment periods. In the comparison of the Gain and Loss, the difference of SUT was 67.41 (SE = 18.45, $p$ < .001) at the first week of treatment, 64.28 (SE = 17.95, $p$ = .001) at the second week of treatment, and 68.69 (SE = 19.63, $p$ = .002) at the last week of treatment. The difference between the Control and the Loss for each treatment period was 90.42 (SE = 18.45, $p$ < .001), 82.60 (SE = 17.95, $p$ < .001), and 98.37 (SE = 19.63, $p$ < .001), respectively. Meanwhile, in the comparison of the Control and Gain, there was no significant difference in SUT for each period. The difference between the Control and the Gain for each treatment period was 23.02 (SE = 18.52, $p$ = .646), 18.32 (SE = 18.02, $p$ = .931), and 29.68 (SE = 19.70, $p$ = .400), respectively.




**(a) Control Group**

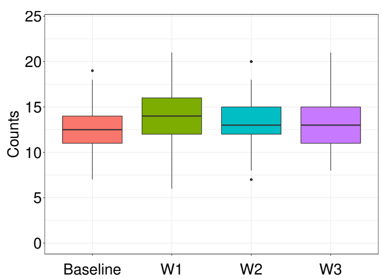



**(b) Gain Group**

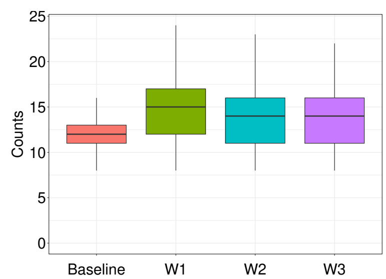



**(c) Loss Group**

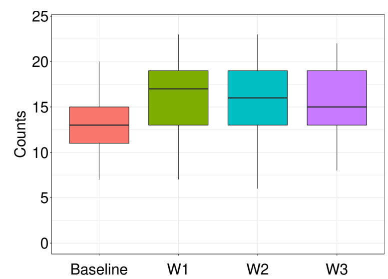
**Figure 9: Daily Average of Regulation Success Counts (Baseline vs. Intervention)**




**(a) Control Group**




**(b) Gain Group**




**(c) Loss Group**

**Figure 10: Daily Average of Regulation Success Counts by Period**

*5.1.2 Daily Average of Regulation Success Counts (RSC).* We also analyzed how RSC was changed over time for each group. Descriptive statistics showed that the daily average of SUT overall increased during the intervention period in all three groups as shown in Figure 9 and 10; Control (Baseline: M = 12.66, SD = 2.47; Intervention W1: M = 13.94, SD = 2.93; Intervention W2: M = 13.52, SD = 2.73; Intervention W3: M = 13.34, SD = 3.04), Gain (Baseline: M = 12.12, SD = 1.85; Intervention W1: M = 14.97, SD = 3.88; Intervention W2: M = 14.25, SD = 3.63; Intervention W3: M = 13.78, SD = 3.18), and Loss (Baseline: M = 12.58, SD = 3.00; Intervention W1: M = 16.48, SD = 3.71; Intervention W2: M = 15.70, SD = 3.95; Intervention W3: M = 15.71, SD = 3.73).

A mixed ANOVA with a Greenhouse-Geisser correction showed that there were significant differences across the four-time periods, F(2.31, 467.92) = 64.98, $p$ < .001, $\eta^2$ = .269, and significant differences between groups, F(2, 202) = 7.12, $p$ < .001, $\eta^2$ = .069, in RSC. There was also a significant interaction effect between period and group, F(4.63, 467.92) = 8.78, $p$ < .001, $\eta^2$ = .069. Following up this interaction indicated that there was no significant difference between groups at baseline and the control group was not significantly changed from 2 weeks of treatment. On the other hand, the mean scores of the other two groups increased at all treatment periods compared with the baseline, each of which is significantly different. In the Gain, the mean difference compared with the baseline was 2.85 (SE = .38, $p$ < .001) at the first week of treatment, 2.13 (SE = .38, $p$ < .001) at the second week of treatment, and 1.66 (SE = .38, $p$ < .001) at the last week of treatment. In the Loss, each mean difference was 3.90 (SE = .38, $p$ < .001), 3.12 (SE = .37, $p$ < .001) and 3.13 (SE = .38, $p$ < .001), respectively.

From the multiple pairwise comparison of groups, we observed that the mean score of Loss was significantly different from those of other two groups over the entire treatment period. In the comparison of the Gain and Loss, the difference of RSC was 1.51 (SE = .60, $p$ = .04) at the first week of treatment, 1.47 (SE = .60, $p$ = .047) at the second week of treatment, and 1.93 (SE = .57, $p$ = .003) at the last week of treatment. The difference between the Control and the Loss for each treatment period was 2.54 (SE = .60, $p$ < .001), 2.18 (SE = .59, $p$ < .001), and 2.37 (SE = .57, $p$ < .001), respectively. In the comparison of the Control and Gain, there was no significant difference in RSC for each period. The difference between the Gain and the Control for each treatment period was 1.03 (SE =.61, $p$ = .272), 0.74 (SE = .60, $p$ = .656), and 0.44 (SE = .57, $p$ = 1.00), respectively.

## 5.2 User Perception and Behaviors for Self-Regulation (RQ3)

To understand how our *GoldenTime* affected users' perception and their actual smartphone usage regulation, we set out the following sub-RQs in relation to well-known processes of self-regulation theories.

- **(RQ3-1)** : What are the users' perceived problematic usage behavior under system-driven timeboxing strategy/design and what are the key factors that contribute to such behavior evaluations?

- **(RQ3-2)** : What are the users' coping strategies under system-driven timeboxing strategy/design?
- **(RQ3-3)** : How does incentive framing affect the user's behavior evaluation and coping strategies?

*5.2.1 RQ3-1: Perceived Problematic Usage and Contributing Factors.* Overall, participants from all three groups (control, gain, and loss) responded that using timeboxing and real-time notification bar design made them reflect upon their habitual phone usage or patterns. Participants also reported this reflection process had become their barometer of self-judgment, as they constantly monitored their current success/failure status based on timeboxing and real-time notification. Here, we categorize two types of problematic behaviors commonly reported by users.

**Habitual unlocking/checking certain apps (n = 126, 61%)**. One commonly reported type of problematic smartphone usage was habitual unlock or checking of certain apps when the usage is not perceived necessary. P21 stated, *"I found myself just unlock and press anything unconsciously, as if in class!"* P2 also reported, *" I would just check my Instagram and Facebook when it was not necessary."*

**Addictive/Unbreakable media usage (n = 121, 59%)**. Severe indulgence in media consumption was another often-reported problem. For example, P13 stated his/her concerns, *"I have this habit... At the library I would decide to take a short break watching YouTube or Facebook videos, but once I get into it, you know, the break takes forever."* P16 also noted, *"Streaming services like YouTube or Netflix... You can't stop once you start!"*

In explanation to the aforementioned behaviors, we found that the following three aspects may have contributed to participants' realization of their problematic smartphone usage—*Distraction*, *Cost* and *Norm*.

**Feeling Distracted (n = 91, 44%)**. Participants perceived their problematic usage when their current task or planned task being postponed or distracted due to smartphone use. Types of distraction include direct distraction and indirect distraction. Direct distraction refers to a situation in which a participant finds difficulty concentrating on his/her current primary task (e.g., studying). For example, P3 noted, *"I was just browsing through Instagram and the alarm would pop up... That's when I realized I was having trouble concentrating."* P22 also stated, *"When the 9-minute alarm pops up, I would then realize I was spending too much time on this."* Indirect distraction refers to a situation in which one's smartphone usage would affect planned or future tasks. P6 said, *"The alarm message that I received before going to bed would remind me of the meeting and all the to-do's of the next day, which helped me to control my usage."* Similarly, P15 reported by saying *"Especially, I liked the alarm message before going to bed. It felt as if the system would slip in and say, 'Hey, you'd better go to bed now. You've got things to do tomorrow!' I liked the fact that it was throwing red flags at me whenever I was using it too much."*

**Realizing Costs and Risks (n = 139 , 68%)**. Smartphone usage incurred costs and risks. Types of perceived costs and risks reported by users were time, health and potential danger. *Time* was the most commonly reported perceived cost. Users generally reflected upon their problematic usage as they observed their time wasted on smartphones. P1 stated, *"As I could see how much money is deducted, I could also find out how much time is wasted. This helped me a lot facing my problem."* P19 also responded, *"I could see how much time I just wasted away by looking at the dashboard."*

Realizing *health issues* due to smartphone overuse (e.g., fatigue, pink eyes) has affected one's judgment. For example, P29 said, *"I usually watch tons of videos before I go to bed, but this popup message kind of alerted me! It felt as if it was telling me, 'You'll get pink eyes again!'. Oh, I also knew my smartphone use was a problem as I always felt 'Oh I'm going to be so tired when I wake up tomorrow'."* P11 stated, *"You see, it is really bad for your eyes if you use your smartphone with the lights off. Oh, and you'll get cross-eyed if you lie on your side and use your smartphone. I knew this all along but would always ignore it anyway.* GoldenTime *made me aware of all these side-effects."*

Participants realized their problematic smartphone usage when they were exposed to *potential danger* due to indulging in phone use (primarily occurs while performing certain behavior or primary tasks). For example, P29 reported, *"I would often check my phone while driving. I know that I'm not supposed to do that. So when the popup message appeared, I thought I should stop checking my phone."* Another similar response from P30 stated, *"I felt like getting a warning as I was walking down the street sending texts. I thought, 'Maybe I should stop and send the rest as I get to that building.'"*

**Violating Norms (n = 31, 15%)**. Users would also perceive their problematic usage when their self-set rule or social norm were violated in their daily lives. Norms can be either *personal* or *social*. For personal norm, P27 said, *"I felt bad that my plans fell through as I watched YouTube clips in a row."* P21 reported, *"I would try to control myself once I realized that I broke my rule, thinking 'Oh, I was not supposed to watch Youtube around this time of the day!"* Social norms would include instances that involved interaction with others. For example, P24 said, *"I think the biggest problem was that I would always use my phone in the dark while my roommate was sleeping..."* Another reported answer from P30 mentioned, *"I felt sorry for people who were looking for empty spots in the cafeteria as I was goofing around with my phone."*

*5.2.2 RQ3-2: Coping Strategies.* To deal with one's perceived problematic smartphone usage, participants would take the following coping strategies. Reported strategies can be largely categorized into following cases: *use* and *non-use*.

**'Use' Strategy (n = 141, 69%)**. 'Use' refers to cases in which participants would come up with self-set disciplines to manage their problematic usage, but still use smartphones. Disciplines include smartphone usage time restriction/regulation or app usage rule/strategy.

Smartphone usage time restriction/regulation refers to one's attempt to control i) total time spent using a smartphone or ii) allocating a specific time slot for smartphone usage. As to usage time restriction/regulation, P23 said, *"I would plan how much time it would take to complete certain tasks and try to fit into that schedule (i.e., 10-minute restriction)."* Another interesting thing to note from the interview was that some participants would deliberately find media consumption that would take less than 10 minutes. For example, P14 said, *"As I tried to restrict my usage time, I would avoid clips that would take longer than 10 minutes."* As to timezone strategy, P2 mentioned, *" I would just quit watching streaming services while I was studying or in the middle of doing something. You know these* *activities are meant to take such a long time, so I would just make time for this. Besides this time, I just made a resolution not to watch these services."* P22 would say, *"I think it sort of created a habit of using my phone only during the certain time period of the day. I thought, 'It's actually quite okay not to use my phone… nothing actually happens!' You know, 10 minutes is more than enough for urgent messages and you can always take a detour, your laptop!"*

Examples of app usage rule/strategy include efforts to restrain the counts of app contents that a participant plans to consume or take only limited use of an app (e.g., only consuming news contents). Efforts to restrain the counts of app contents would include following examples: *"I would decide how many video clips I would watch before going to bed and turn off my phone right away!* (P6), and *"Binge-watching cartoons... that takes forever. Considering that I had to watch until episode 100, I would set 10, 20 episodes a day.* (P21). Selective use of an app would include the following instances: *"I tried to use only English dictionary and Google search during class"* (P9) and *"I used to check every single status update of my friends and send texts… but I decided not to. I told myself that I would only text back to urgent messages during class"* (P12).

**'Non-Use' Strategy (n = 94, 46%)**. Non-Use strategy, as its name indicates, aims for a complete non-use of a phone. The strategy includes 'Preventive/Block method' and 'Physical Non-Use'. Preventive/block method includes refusing to receive notification of certain apps or turning off the smartphone. For example, P13 mentioned, *"I would just turn off my Facebook notifications before I go to class, otherwise I knew I would just use my phone.* P14 said, *"In the case of non-use, I would keep using it when it was urgent or there were things I really needed to say... But most of the time I would turn off my notifications.* Participants also resorted to 'Physical non-use' strategy by placing their smartphones distant from their current locations. *"I would just leave my phone or have it charged somewhere else when I was eating"*, said P1. Another strategy included turning to an alternative to restrain smartphone usage (e.g., exercise). *"I think this really helped me stop using my smartphone when it was not necessary. I would tell myself like, 'Let's just do something else.. Maybe I could read books, study harder or do some exercise.' I think I tried to act more consciously."* (P1)

Despite the aforementioned positive effects, some participants attempted a detour (e.g., using a laptop or other devices) to avoid using *GoldenTime*. For example, P23 noted, *"I would use PC chat instead of using smartphones."*

*5.2.3 RQ3-3: Evaluation and Coping Process.* We compared the overall process of behavior evaluation and coping among different groups. We found different patterns across groups: (1) control group with the demotivation process, (2) gain group with the devaluation process, (3) loss group with the revaluation process.

**Control Group: "Demotivation Process"** In the first half of the intervention period, the experience of success or failure in self-regulation helped participants to self-regulate usage behaviors. However, as time passed, its effect waned. The major reason was that the motivation of self-regulation had weakened. Participants often mentioned that it was because there were no rewards or penalties whatsoever regarding the behavior of either successful or failed self-regulation. One participant said, *"as it was noted as a failure in self-regulation, I guess there was a bit of a psychologically suppressive effect of so. Surely, as I kept using it since supposedly no penalties were imposed, my will weakened and the resolution for self-regulation slightly loosened."* (P21).

Another reason was that their response toward intervention had turned dull. They stated that as they got used to the intervention process that lasted repetitively, the response to this had turned dull, and ended up ignoring the intervention. One participant said, *"as time went on, the notification alarm turned dull to the point of exceeding 10 minutes so it crossed my mind that even if I fail, it comes up again."* (P29).

As we approached the latter half of the intervention, the coping strategies also loosened up a bit more, in comparison with the former half. In the first half, participants tended to set up specific usage plans and strategies focused on the usage time or success/failure frequency. One participant said, *"in the first half of the experiment, there were tons of stimuli while viewing the usage time or the number of successful self-regulation. As I hit upon the idea that I should do a bit better today than yesterday or tomorrow than today and so forth, I clearly set and abided by the usage time range but, I guess, I lost tension. Afterward, I degenerated, thinking since I withstood it five times. That is, I succeeded in self-regulation five times, no bones are broken. It's okay."* (P24).

**Gain Group: "Devaluation Process"** The prevailing responses related to the monetary value concerning successful timeboxing were represented by a sense of achievement and self-efficacy. Participants mentioned that the sense of achievement due to the successful self-regulation and self-efficacy (i.e., a positive feeling that an incentive was obtained by endeavoring for self-regulation) boosted their self-esteem. One participant mentioned, *"It was great to feel a sense of accomplishment and being rewarded. It was brought home to me that time is money. As I abstain myself, I make a name in the world and mint money."* (P12).

The psychological process that influenced the perceived monetary value related to failed self-regulated behavior was the regret of missing or wasting golds. However, this process manifested a more complex aspect, showing two opposite reactions to failures. Those users, who embraced the sense of regret towards and feeling of having wasted the missed gold as part of self-reflection, were awakened to the price one has to pay concerning failure in self-regulation and highly evaluated the monetary value of failures. One participant mentioned, *"I missed gold... [I] felt reluctant to give up on it. I felt it was a waste, as I hooked up the wasted time with bucks. I flashed through my mind I should do better next time"* (P14). In contrast, the users, who embraced the missed gold as a 'one-time opportunity missed in an infinite stream of rewards, showed a tendency of self-rationalization towards failed self-regulation behavior. As a result, the monetary value associated with phone usage is devalued over time. One participant said, *"although I flunked in self-regulation one day since I can earn again by doing self-regulation in the following day and the lost amount is negligible. This much is no big deal."* (P16).

The gain group users showed a tendency to set up plans and strategies of self-regulation usage behavior to earn gold. They mainly set a gold earning goal on a daily basis and freely regulated usage within the possible range of achieving the goal. One participant said, *"You earn money, using the app. So it brings an objective. ... Let me earn at least this much"*(P13). Another participant said, *"I resolved to fill 10 golds [in the notification bar gold image] every single day."* (P19). As time passed by, we observed two different attitudes on the monetary value related to the failed self-regulation behaviors, which had a significant influence on coping strategies. Through failure in self-regulation, those users who realized the cost of or price one must pay for the problematic usage behavior strived for self-regulation by raising the goals (e.g., raising the gold earning goal) to make up for lost golds. One participant said, *"while viewing the daily notification, since I've earned this much yesterday, how about earning a bit more today? I set and jot down the goals likewise"* (P13). In contrast, those who perceived the incentive mechanism as "an infinite opportunity" for gold mining devalued the cost of failures like those users in the control group; and thus, their coping strategies gradually waned. One participant said, *"feasting my eyes on the amassed gold in the meantime, I got slightly loose later on. Since I could earn golds later despite failures, I came to use my phone more resiliently and set goals in an easygoing fashion like that"* (P18). Another participant stated, *"towards the latter half, my will and enthusiasm declined and got insensible towards [warning] notifications. My need for self-regulation died down since I didn't feel the amount of gold that was being piled up wasn't big against not using the phone by self-regulation."* (P18)

**Loss Group: "Revaluation Process"** Loss framing gradually formed a mental model of "usage fee," meaning that successful self-regulation resulted in guarding pre-assigned monetary rewards, and this also brought a sense of relief and self-efficacy. This formation process guided participants to reevaluate the monetary value associated with each timebox, and this led them to be self-conscious about their usage behaviors and plan more elaborate coping strategies for self-regulation. Users gradually embraced the sense of loss toward the gold deducted, and self-reflection regarding the problematic usage behavior helped them to realize the cost of self-regulation failures. This process guided our participants to *revaluate* the monetary value of a timebox.

Participants learned a sense of loss with " deducted gold." One participant said, *"as the deduction kept repeating, at some point I thought every single usage behavior all entailed a price that had to be paid. Just like many a little makes a mickle... That's why I practiced further caution/prudence in usage behavior."* (P7) Another participant commented, *"It was an amazing experience. I usually won't be able to pick up a 50-won coin [approx. a nickel] fallen on the street; however, now that I was about to take off 50-won in exchange for using the app. Mindfully, I exercised self-regulation. And overtime, the deducted amount kept accumulating, and I realized the price I had to pay was greater than I thought."* (P19). In addition, successful self-regulation brought relief as one participant said, *"when the gold icon (on the notification bar) was full, I felt a sense of relief. I felt something like I pulled it off. I fought well to save it"* (P5).

In general, the loss group tended to set up plans and strategies of self-regulation behavior centered on the amount of usage time. They usually set the amount of time of smartphone usage on a daily basis and funneled efforts into self-regulation, to avoid exceeding the targeted amount of time. Such a behavior strategy originates from a psychological mechanism of loss aversion. A user's loss aversion was expressed as a tendency of setting up rigorous coping strategies over time in order to minimize gold deduction. Participants strived to reduce the unnecessary usage by planning in detail what app to use for how long and in what time range a prior, to the actual usage and to maintain the gold and self-regulation behavior by minimum necessary usage. One participant mentioned, *"By and large, I've used it, being super discreet and heedful. I set plans for the timeline and I set the number of minutes to use for certain apps. I developed the notion of 'binding together for usage' to save time. For instance, when doing Facebook, I also log on Instagram to skim through and handle piled text messages in the bundle"* (P7).

# 6 DISCUSSION

We discuss the effectiveness of the proposed system-driven timeboxing and micro-financial incentive mechanism to regulate smartphone usage. In addition, we provide practical design guidelines for persuasive technology design related to discouraging undesired behaviors.

## 6.1 Supporting Self-Regulated Usage Behavior

Our experimental results showed that system-driven timeboxing effectively supports self-regulated usage behavior. The system-driven timeboxing helped users to consciously use smartphones with purposes and goals within a timebox. In terms of time management, it served as a time management tool that helped users control and manage their usage time. Furthermore, it supported users not to lose their goal of regulated usage behavior by allowing limited usage time so that they can cope with various usage contexts. In comparison with the existing just-in-time intervention methods such as LocknType which induces interaction friction on each screen on activity [34], our approach relaxes restraint intensity (or increases flexibility) in that a warning is given one an hourly basis. The system-driven timeboxing also helped users to maintain regulated usage behaviors and form good behaviors such as "time planning." Existing persuasive system design studies emphasized the importance of forming "good habits" in order to maintain sustainable and long-term behavioral changes. A prior study [50] reviewed various intervention approaches to help mitigate problematic usage behaviors, but experimental studies on scaffolding good habits were relatively limited. Our work showed that system-driven timeboxing as a novel design dimension offers new ways of guiding self-regulated behaviors and learning positive usage behaviors.

Our results also showed that proactive warning alarms in the system-driven timeboxing effectively supported the self-regulation process for smartphone usage. The proactive warning alarm helped users to continuously observe and recognize their problematic usage behaviors in various usage contexts. This result aligns with the implications of existing intervention studies [38] such as when unlocking phones or arriving at specific places. Timely warning with push notifications greatly improves self-monitoring and this will help users to sustain target behaviors (e.g., limiting use). In this respect, *GoldenTime*'s proactive warning alarms were effective in self-tracking and regulating their usage behaviors. For example, our experimental results showed that the proactive warning alarm served as a "reminder" to return to the main task from problematic usage behavior. According to the PRIME theory for behavior changes [8], a user's plans as self-conscious intentions are critical, and there is a strong need for an intervention design that timely reminds such plans for positive behavioral change. Our experimental results showed that the proactive warning alarm supported the main task reminder can be aligned with the intervention design guidelines of PRIME theory.

## 6.2 Effects of Incentive Framing on System-Driven Timeboxing

Overall, our results showed that micro-financial incentives were effective in reinforcing positive behavioral change under the system-driven timeboxing. Incentive groups showed positive effects on regulating smartphone usage, but the control group did not show improvement in maintaining the regulated usage behavior over time due to the lack of behavioral reinforcers. Our experimental results supported our hypothesis that loss framing would be more effective in inducing regulated usage behavior than gain framing; this result is consistent with prior studies in other domains [59]. The effectiveness of loss framing attributes to the endowment effect and loss aversion phenomenon according to behavioral economics. Indeed, the loss group frequently mentioned expressions related to the "deduction of my money" (n = 47, 68%). As a result, the endowment effect and loss aversion helped them to have a *revaluation process* on the monetary value of the smartphone usage behavior. Furthermore, we note that prior studies on financial incentives mainly paid incentives less frequently (e.g., at the end of the experiment or on a daily basis), and incentive tracking is less explicit. Unlike prior studies, our approach has "recurring" timeboxing every hour, and this creates an opportunity for rewarding users with a micro-incentive in different framing for usage regulation.

None of the prior studies systematically explored how users react to "recurring" financial incentives over time, and our results showed that there were clear differences between gain and loss framing groups (i.e., devaluation vs. revaluation). At the beginning of the intervention period, the loss group felt negative emotions (e.g., loss) through the experience of deducting incentives. However, this experience was regarded as a "positive loss" that perceived self-reflection and awareness of problematic usage behavior in the micro-incentive mechanism. In addition, through these experiences, they had the concept of usage fee for smartphone usage behavior and tried to adapt to micro-incentive mechanisms by making specific plans or strategies for usage time and actions. As a result, the loss mechanism formed a mental model of "metering (provides real-time deduction amount for smartphone usage)" to users who experience the mechanism, thereby making them revaluate the monetary value of phone use. In contrast, the gain mechanism made users feel positive (e.g., achievement) through a regulation success, thereby reinforcing regulated usage behavior. However, the experience of using the gain mechanism also included the devaluation tendency with micro-incentives. This may be the result of the self-rationalization of degrading the monetary value of regulated usage behavior by treating it as puny money for intentional ignoring. These psychological factors were heightened with unintended success experiences (e.g., the experience of earning golds without any efforts such as when using different devices). In addition, the micro-incentive mechanism was recognized as an "infinite opportunity" as time passed by. Such user perception of a micro-incentive could be interpreted as "house money," which refers to a money that is easily obtained, thus incurring low perceived value in users [60].

Different experiences (i.e., revaluation vs. devaluation) of micro-incentive mechanisms can be interpreted through Behavioral economics theory. First, different trends were interpreted as the result of being influenced by the "reference dependency" based on an existing theory [29]. In short, the reference points of the two mechanisms for incentives were different. The behavioral economics theory explained that the reference point is the standard point for determining loss or benefit. In our design, since incentives are prepaid in the Loss mechanism, there is no experience of incentive gains (+) in the subsequent intervention process, and only deduction (-) experiences exist in the long term. In our study, since the Loss mechanism prepays incentives, users do not experience incentive gains (+) in the subsequent intervention process, and only experience deductions (-) in the long term. On the other hand, in the Gain mechanism, since the reference point starts at 0 (no advance payment), even if the regulation success or failure experiences are repeated, only the gain (+) experience exists in the long term. This can be explained through the editing rules of prior study. Editing rules describe the tendency of combining gain and loss experiences to determine profit or loss (i.e., a loss after having gain is less painful). Overall, we argue that loss framing is more effective in micro-incentive based behavioral reinforcement. In particular, through our experiments, we obtained implications that such editing rules can be further deepened in the process of *continuous decision-making*, such as recurring timeboxing, and can affect behavioral changes in a long-term perspective.

## 6.3 Design Implications

Our results provide practical design guidelines for persuasive technology design related to discouraging undesired behaviors.

*6.3.1 Towards Flexible and Context-Aware Proactive Intervention.* We explored the design of a timeboxing-based proactive intervention to induce behavioral changes and experimentally showed promising evidence on behavioral changes. However, our work was not able to fully explore various contexts and situations of the users for digital wellbeing [44]. Prior studies demonstrated the usefulness of a context-aware proactive intervention such as location [31] and app usage [46]. Current timeboxing mechanisms can be extended to enable "context-aware" adaptive interventions. As mentioned by our participants, timeboxing can be skipped if a user is using a mobile navigation app for long-distance driving. Users can set up such adaptive behaviors via rule-based specifications. Alternatively, advanced sensing can be used to enable context-triggered actions; e.g., upon sensing a user is driving, proactive timeboxing is automatically disabled. Depending on what kinds of context sensing are feasible, a timeboxing-based proactive intervention can incorporate fine-grained control on its detailed operations. Context-awareness provides "flexibility" in proactive intervention, but one critical dimension is to consider to what degree flexibility will be permitted. This kind of flexibility support can even consider introducing the concept of margin in the goal evaluation criteria (e.g., giving 30 seconds of margin beyond the threshold) [27]. Furthermore, context-aware operations need to be clearly specified (by either users or other persons such as parents). As in traditional context-aware systems design, end-user programming methods such as trigger-action programming (if-this-then-that) and block-based programming must be supported [30].

*6.3.2 Exploring Micro-incentive Design for Timeboxing.* Our study showed the feasibility of leveraging micro-incentives for behavior changes. In our micro-incentive design, we set the self-regulated smartphone usage behavior as the target behavior. Our design experience can be also applied to other just-in-time intervention applications that aim to intervene in fine-grained user behaviors (e.g., sedentary or smoking instances). This work laid the foundation for exploring the extended design space for micro-incentive mechanisms with system-driven timeboxing. Several design dimensions to explore include 1) reward target behavior, 2) reward goal, and 3) reward methods. First, the reward target as a direct object of the reward may be period (i.e., time) or event based (e.g., situations). Second, the reward goal is the design of a user's goal-setting in terms of earned reward amount (e.g., I wanted to earn 10 USD), which could be short- or long-term goals. Third, the reward methods could be 1) setting reward amount (fixed or variable amount), 2) framing strategies (gain vs. loss), and 3) reward tracking and payment methods (real-time transfer of fiat money, or blockchain-based virtual currency). Specifically, for variable reward selection, we can extend the design method in which the amount of money is adaptively set based on a user's previous behaviors (e.g., continuous success or failure). Reward methods can also consider social aspects; e.g., when parents are using the services for training children's smartphone usage. It is also possible to set up a commitment device that may transfer committed stakes to the friends upon goal failures [42, 43].

*6.3.3 Data-Driven Actionable Insights for Better Usage Planning.* The experimental results explained that the loss mechanism formed *mental models* of "usage metering" or "pay as you go" for smartphone usage behavior, thereby helping them to sustain regulated usage behavior via coping strategies. From the perspective of persuasive technology design, we emphasize the need to provide a user interface that can effectively support self-regulation in accordance with their mental model. This can be supported through data analytics and visualization. One approach of data analytics is to analyze individual contextual data that are closely related to behavioral changes (e.g., under what conditions did a user often fail missions?) A prior study analyzed individual context information (e.g., sleep, mood) that has a sufficient correlation with well-being to lead better self-reflection and behavior change [6]. Data analytics can be extended beyond an individual level. Community-wide behavioral patterns could be learned by analyzing the interaction log data [39]; (e.g., under what conditions did people often succeed in self-regulation?) Individual- and community-level insights must be properly verbalized or visualized for effective delivery to the users [6, 39].

## 7 LIMITATION AND FUTURE WORK

There are several limitations in this study that could be addressed in future works. First, our intervention design did not consider the user preference for timeboxing setting, which potentially increased false-positive notifications. In this study, we rather tried to use uniform parameters to fairly evaluate the proposed intervention. Although it is useful to provide personalized settings in that the usage time plan may different person to person, we excluded this option to avoid confounding effects: incentive amounts were closely related to timeboxing configuration. For a better experience of using timeboxing, future work needs to consider using context-aware methods or experience sampling method similar to Lukoff et al.'s work [49]. This may deepen our analysis of understanding how engagement with micro-incentive mechanisms change according to smartphone usage instances.

Second, there was no follow-up period in our user study design. The current study was sufficient to compare and observe short-term effects while using the intervention app. However, it may not be sufficient for observing the long-term effects of self-regulation or for an in-depth analysis of experience with micro-incentive mechanisms. It is still important to observe how users' self-regulated behavior changes in the absence of incentives after experiencing the mechanism. One of the most common findings of incentives is a positive effect in the short-term but crowding out in the long-term [21]. Future work should consider long-term experimental design including follow-up period.

Third, our findings should be carefully understood because most of the participants were university students. In addition, the sample was limited to Android users, which does not reflect the general population. For generalizability of the findings, follow-up studies needs to include users of various age groups and mobile platforms.

## 8 CONCLUSION

We designed and implemented *GoldenTime*, a smartphone intervention app to support continuous self-tracking and regulation on smartphone use. The *GoldenTime*, using micro-financial incentives under system-driven timeboxing was designed based on behavioral economics, self-regulation theory, and the results of prior empirical studies on smartphone regulation. Our between-subject study (n = 210) over the four-week demonstrated that compared with the baseline, users who used the micro-incentive mechanism experienced regulation effect in terms of usage time. The comparative analysis between groups showed that micro-financial incentives, when framed as loss, were more effective than the other cases over time. Furthermore, our qualitative analysis revealed that the overall process of behavior evaluation and coping differs significantly, ranging from the demotivation process to the devaluation and revaluation process. Our findings provided several important design implications for self-regulation. We believe that our approach can be extended to a variety of behavioral change domains such as weight loss that should continuously promote self-regulation.

## ACKNOWLEDGMENTS

This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2017M3C4A7083534) and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2020R1F1A1074161)

# REFERENCES


[1] Elena Agapie, Daniel Avrahami, and Jennifer Marlow. 2016. Staying the course: System-driven lapse management for supporting behavior change. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1601 Broadway, Times Square, New York City., 1072–1083.

[2] Nabil Alshurafa, Jayalakshmi Jain, Rawan Alharbi, Gleb Iakovlev, Bonnie Spring, and Angela Pfammatter. 2018. Is more always better?: discovering incentivized mHealth intervention engagement related to health behavior trends. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 153.

[3] 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*. ACM, 1601 Broadway, Times Square, New York City, 95–106.

[4] Albert Bandura et al. 1991. Social cognitive theory of self-regulation. *Organizational behavior and human decision processes* 50, 2 (1991), 248–287.

[5] Steve Benford, Chris Greenhalgh, Gabriella Giannachi, Brendan Walker, Joe Marshall, and Tom Rodden. 2012. Uncomfortable interactions. In *Proceedings of the sigchi conference on human factors in computing systems*. ACM, 1601 Broadway, Times Square, New York City, 2005–2014.

[6] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 5 (2013), 1–27.

[7] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[8] Ross Buck. 1985. Prime theory: An integrated view of motivation and emotion. *Psychological review* 92, 3 (1985), 389.

[9] James Clawson, Jessica A Pater, Andrew D Miller, Elizabeth D Mynatt, and Lena Mamykina. 2015. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, ACM, 1601 Broadway, Times Square, New York City, 647–658.

[10] Emily IM Collins, Anna L Cox, Jon Bird, and Daniel Harrison. 2014. Social networking use and RescueTime: the issue of engagement. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 1601 Broadway, Times Square, New York City, 687–690.

[11] Sunny Consolvo, David W McDonald, and James A Landay. 2009. Theory-driven design strategies for technologies that support behavior change in everyday life. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, ACM, 1601 Broadway, Times Square, New York City, 405–414.

[12] Sunny Consolvo and Miriam Walker. 2003. Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing* 2, 2 (2003), 24–31.

[13] Anna L Cox, Sandy JJ Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 1601 Broadway, Times Square, New York City, 1389–1397.

[14] 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*. ACM, 1601 Broadway, Times Square, New York City, 179–194.

[15] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191.

[16] Stephen Ray Flora. 2012. *Power of Reinforcement, The*. SUNY Press, Albany, NY, United States.

[17] Brian J Fogg. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002, December (2002), 2.

[18] Brian J Fogg. 2019. *Tiny Habits: The Small Changes That Change Everything*. Houghton Mifflin Harcourt, 125 High St, Boston, MA 02110.

[19] David Gibson, Nathaniel Ostashewski, Kim Flintoff, Sheryl Grant, and Erin Knight. 2015. Digital badges in education. *Education and Information Technologies* 20, 2 (2015), 403–410.

[20] Fausto Giunchiglia, Mattia Zeni, Elisa Gobbi, Enrico Bignotti, and Ivano Bison. 2018. Mobile social media usage and academic performance. *Computers in Human Behavior* 82 (2018), 177–185.

[21] 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.

[22] Deborah J Hennrikus and Robert W Jeffery. 1996. Worksite intervention for weight control: a review of the literature. *American Journal of Health Promotion* 10, 6 (1996), 471–498.

[23] Stephen T Higgins, Kenneth Silverman, and Sarah H Heil. 2007. *Contingency management in substance abuse treatment*. Guilford Press, 370 Seventh Avenue Suite 1200 New York, NY 10001-1020.

[24] 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*. ACM, 1601 Broadway, Times Square, New York City, 4746–4757.

[25] Alexis Hiniker, Bongshin Lee, Kiley Sobel, and Eun Kyoung Choe. 2017. Plan & play: supporting intentional media use in early childhood. In *Proceedings of the 2017 Conference on Interaction Design and Children*. ACM, 1601 Broadway, Times Square, New York City, 85–95.

[26] Robert W Jeffery, Jean L Forster, Simone A French, Steven H Kelder, Harry A Lando, Paul G McGovern, David R Jacobs Jr, and Judith E Baxter. 1993. The Healthy Worker Project: a work-site intervention for weight control and smoking cessation. *American journal of public health* 83, 3 (1993), 395–401.

[27] Gyuwon Jung, Jio Oh, Youjin Jung, Juho Sun, Ha-Kyung Kong, and Uichin Lee. 2021. "Good Enough!": Flexible Goal Achievement with Margin-based Outcome Evaluation. In *In CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA.

[28] Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. 1991. Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives* 5, 1 (1991), 193–206.

[29] 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, 5 Toh Tuck Link Singapore, 596224. Singapore., 99–127.

[30] Inyeop Kim, Hwarang Goh, Nematjon Narziev, Youngtae Noh, and Uichin Lee. 2020. Understanding User Contexts and Coping Strategies for Context-Aware Phone Distraction Management System Design. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 134 (Dec. 2020), 33 pages. https://doi.org/10.1145/3432213

[31] Inyeop Kim, Gyuwon Jung, Hayoung Jung, Minsam Ko, and Uichin Lee. 2017. Let's FOCUS: mitigating mobile phone use in college classrooms. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–29.

[32] Jaejeung Kim, Chiwoo Cho, and Uichin Lee. 2017. Technology supported behavior restriction for mitigating self-interruptions in multi-device environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–21.

[33] 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.

[34] 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*. ACM, 1601 Broadway, Times Square, New York City, 1–12.

[35] Mijung Kim. 2014. *The effects of external cues on media habit and use: Push notification alerts and mobile application usage habits*. Michigan State University, 220 Trowbridge Rd East Lansing, MI 48824, USA.

[36] Minsam Ko, Seungwoo Choi, Subin Yang, Joonwon Lee, and Uichin Lee. 2015. FamiLync: facilitating participatory parental mediation of adolescents' smartphone use. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1601 Broadway, Times Square, New York City, 867–878.

[37] 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*. ACM, 1601 Broadway, Times Square, New York City, 998–1010.

[38] 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*. ACM, 1601 Broadway, Times Square, New York City, 1235–1245.

[39] Nicholas D Lane, Li Pengyu, Lin Zhou, and Feng Zhao. 2014. Connecting personal-scale sensing and networked community behavior to infer human activities. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1601 Broadway, Times Square, New York City, 595–606.

[40] Robert LaRose. 2007. Uses and gratifications of Internet addiction. In *Internet addiction: A handbook and guide to evaluation and treatment*. Wiley Online Library, Hoboken, New Jersey, Chapter 4, 55–72.

[41] Robert LaRose, Carolyn A Lin, and Matthew S Eastin. 2003. Unregulated Internet usage: Addiction, habit, or deficient self-regulation? *Media Psychology* 5, 3 (2003), 225–253.

[42] Hyunsoo Lee, Auk Kim, Hwajung Hong, and Uichin Lee. 2021. Sticky Goals: Understanding Goal Commitments for Behavioral Changes in the Wild. In *In CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA.

[43] Hyunsoo Lee, Uichin Lee, and Hwajung Hong. 2019. Commitment devices in online behavior change support systems. In *Proceedings of Asian CHI Symposium*

*2019: Emerging HCI Research Collection*. ACM, 1601 Broadway, Times Square, New York City, 105–113.
[44] Uichin Lee, Kyungsik Han, Hyunsung Cho, Kyong-Mee Chung, Hwajung Hong, Sung-Ju Lee, Youngtae Noh, Sooyoung Park, and John M. Carroll. 2019. Intelligent positive computing with mobile, wearable, and IoT devices: Literature review and research directions. *Ad Hoc Networks* 83 (2019), 8 – 24.
[45] Brian Y Lim, Judy Kay, and Weilong Liu. 2019. How Does a Nation Walk?: Interpreting Large-Scale Step Count Activity with Weekly Streak Patterns. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 57.
[46] Markus Löchtefeld, Matthias Böhmer, and Lyubomir Ganev. 2013. AppDetox: helping users with mobile app addiction. In *Proceedings of the 12th international conference on mobile and ubiquitous multimedia*. ACM, 1601 Broadway, Times Square, New York City, 1–2.
[47] Edwin A Locke and Gary P Latham. 2006. New directions in goal-setting theory. *Current directions in psychological science* 15, 5 (2006), 265–268.
[48] Danielle Lottridge, Eli Marschner, Ellen Wang, Maria Romanovsky, and Clifford Nass. 2012. Browser design impacts multitasking. In *Proceedings of the human factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, SAGE Publications, Newbury Park, California., 1957–1961.
[49] 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.
[50] Ulrik Lyngs, Kai Lukoff, Petr Slovak, Reuben Binns, Adam Slack, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. 2019. Self-Control in Cyberspace: Applying Dual Systems Theory to a Review of Digital Self-Control Tools. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 1601 Broadway, Times Square, New York City, 1–18.
[51] Brian E Mavis and Bertram E Stöffelmayr. 1994. Multidimensional evaluation of monetary incentive strategies for weight control. *The Psychological Record* 44, 2 (1994), 239–252.
[52] Marc S Mitchell and Paul I Oh. 2016. Framing financial incentives to increase physical activity among overweight and obese adults. *Annals of internal medicine* 165, 8 (2016), 599–600.
[53] Tanushree Mitra, Clayton J Hutto, and Eric Gilbert. 2015. Comparing person-and process-centric strategies for obtaining quality data on amazon mechanical turk. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, ACM, 1601 Broadway, Times Square, New York City, 1345–1354.
[54] David C Mohr, Stephen M Schueller, Enid Montague, Michelle Nicole Burns, and Parisa Rashidi. 2014. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. *Journal of medical Internet research* 16, 6 (2014), e146.
[55] 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*. ACM, 1601 Broadway, Times Square, New York City, 435–444.
[56] Harri Oinas-Kukkonen, Khin Than Win, Evangelos Karapanos, Pasi Karppinen, and Eleni Kyza. 2019. *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems: 14th International Conference, PERSUASIVE 2019, Limassol, Cyprus, April 9–11, 2019, Proceedings*. Vol. 11433. Springer, 11 West 42nd Street[6] in Manhattan, New York City.
[57] 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*. ACM, 1601 Broadway, Times Square, New York City, 1–12.
[58] Viktoria Pammer and Marina Bratic. 2013. Surprise, surprise: activity log based time analytics for time management. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1601 Broadway, Times Square, New York City, 211–216.
[59] 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.
[60] Jiaxi Peng, Danmin Miao, and Wei Xiao. 2013. Why are gainers more risk seeking. *Judgment and Decision Making* 8, 2 (2013), 150.
[61] 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.
[62] Sasank Reddy, Deborah Estrin, Mark Hansen, and Mani Srivastava. 2010. Examining micro-payments for participatory sensing data collections. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 1601 Broadway, Times Square, New York City, 33–36.
[63] Paul Romanowich and RJ Lamb. 2013. The effect of framing incentives as either losses or gains with contingency management for smoking cessation. *Addictive behaviors* 38, 4 (2013), 2084–2088.
[64] 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*. ACM, 1601 Broadway, Times Square, New York City, 284–296.
[65] Mical Kay Shilts, Marcel Horowitz, and Marilyn S Townsend. 2004. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. *American Journal of Health Promotion* 19, 2 (2004), 81–93.
[66] Burrhus F Skinner. 1963. Operant behavior. *American psychologist* 18, 8 (1963), 503.
[67] Vincent W-S Tseng, Matthew L Lee, Laurent Denoue, and Daniel Avrahami. 2019. Overcoming Distractions during Transitions from Break to Work using a Conversational Website-Blocking System. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 1601 Broadway, Times Square, New York City, 1–13.
[68] Alexander JAM Van Deursen, Colin L Bolle, Sabrina M Hegner, and Piet AM Kommers. 2015. Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. *Computers in human behavior* 45 (2015), 411–420.
[69] Kevin G Volpp, Andrea B Troxel, Mark V Pauly, Henry A Glick, Andrea Puig, David A Asch, Robert Galvin, Jingsan Zhu, Fei Wan, Jill DeGuzman, et al. 2009. A randomized, controlled trial of financial incentives for smoking cessation. *N Engl J Med* 360 (2009), 699–709.
[70] Tetsuo Yamabe, Vili Lehdonvirta, Hitoshi Ito, Hayuru Soma, Hiroaki Kimura, and Tatsuo Nakajima. 2009. Applying pervasive technologies to create economic incentives that alter consumer behavior. In *Proceedings of the 11th international conference on Ubiquitous computing*. ACM, 1601 Broadway, Times Square, New York City, 175–184.