GoalKeeper: Exploring Interaction Lockout Mechanisms for Regulating Smartphone Use

JAEJEUNG KIM, KAIST, Republic of Korea
HAYOUNG JUNG, KAIST, Republic of Korea
MINSAM KO, Hanyang University, Republic of Korea
UICHIN LEE*, KAIST, Republic of Korea

Many people often experience difficulties in achieving behavioral goals related to smartphone use. Most of prior studies approached this problem with various behavior change strategies such as self-reflection and social support. However, little is known about the effectiveness and user experiences of restrictive and coercive interventions such as blocking. In this work, we developed "GoalKeeper," a smartphone intervention app that locks the user into the self-defined daily use time limit with restrictive intervention mechanisms. We conducted a four-week field experiment with 36 participants to investigate the effects and user experiences of varying intensities of restrictive interventions. The results showed that restrictive mechanisms are more effective than non-restrictive mechanisms such as warning. However, we found that restrictive mechanisms caused more frustration and pressure to the users, mainly due to diversity of usage contexts and needs. Based on our study results, we extracted practical implications for designing restrictive mechanisms that balance the intervention effectiveness for behavioral changes and the flexibility for user acceptability.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile devices**; *Empirical studies in HCI*; *Empirical studies in ubiquitous and mobile computing*.

Additional Key Words and Phrases: interaction lockout, lockout mechanism, lockout intensity, smartphone non-use, restrictive technology, commitment device, self-imposed restriction, persuasive technology, behavior change, ILM

ACM Reference Format:
Jaejeung Kim, Hayoung Jung, Minsam Ko, and Uichin Lee. 2019. GoalKeeper: Exploring Interaction Lockout Mechanisms for Regulating Smartphone Use. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 16 (March 2019), 29 pages. https://doi.org/10.1145/3314403

16

1 INTRODUCTION

People set many goals, but they often fail to achieve them. One explanation of this discrepancy can be devised using the concept of the conflicting two selves [21]. The rational "long-run self" initially sets a goal, but the irrational, impulsive "short-run self" pursues immediate gratification afterwards, leading to the failure to achieve

*This is the corresponding author

Authors' addresses: Jaejeung Kim, KAIST, Daejeon, Republic of Korea, jjk@kaist.ac.kr; Hayoung Jung, KAIST, Daejeon, Republic of Korea, hayoung.jung@kaist.ac.kr; Minsam Ko, Hanyang University, Seoul, Republic of Korea, minsam@hanyang.ac.kr; Uichin Lee, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea, uclee@kaist.ac.kr.

2474-9567/2019/3-ART16 $15.00
https://doi.org/10.1145/3314403

the goal. For example, in the context of smartphone use, we frequently observe this type of instant gratification-seeking behavior, such as checking emails, using social media, and playing mobile games due to the ubiquity and availability of these devices.

In this work, we consider smartphone usage mediation as a focusing lens for assessing how we design interactive technologies to bridge this gap. Smartphone overuse is an important problem owing to its various negative effects on physical [28] and mental health [52], safety [29, 35], productivity [31, 33], social relationships [37, 41, 67] and many more areas.

The majority of prior studies have tackled this problem with behavior change strategies such as self-tracking /reflection to increase self-awareness and/or notification/dialog support to appraise and recommend positive behaviors[15, 26, 30, 38]. In addition, recent studies have explored various restrictive approaches such as selective and complete smartphone blocking, and have documented some examples of positive effects (e.g., reducing distraction and increasing productivity) [26, 31]. There are numerous commercial applications that leverage such self-limiting strategies [6, 46, 62] as well. Furthermore, there recent efforts of governments to apply restrictive and coercive methods of managing smartphone use in safety-critical situations such as walking and driving [3, 11]. For example, Apple's Do Not Disturb While Driving Mode automatically locks out a driver's phone while driving by following the NHTSA guidelines.

In this article, we use the term *interaction lockout mechanism (ILM)* to describe a technology-based intervention that restricts or degrades the interactivity of a device to influence interaction patterns or behaviors. Lockout is a term coined by Boehm *et al.* [12] to represent the time interval between system feedback and the point at which the system is ready for the subsequent interaction. During this interval, the system may restrict user inputs or ignore inputs made by the user. We designed intervention mechanisms leveraging lockout, by varying the lockout duration patterns. Example ILMs include locking up smartphones, blocking apps/websites, disabling touch input, delaying the display of content on the screen, or even disabling the Internet connection. This approach is more direct and enforcing than well-known interventions such as self-tracking and dialog support. One may argue that the interaction lockout strategy may be unethical in some ways due to its restrictive characteristics. However, in the field of behavioral economics, the *voluntary use* of restrictive mechanisms has been widely used as a motivating strategy for goal achievement (known as commitment devices or contracts). To be precise, a commitment device is defined as "an attempt to enforce people's *voluntarily imposed restrictions* until they have accomplished their goals, or their *voluntarily imposed penalties* for failing to accomplish their goals" [63].

Despite the abundance of commercial apps supporting *enforced self-regulation* (e.g., disabling some part or all smartphone functionalities according to the user's self-defined goal), there is a lack of systematic studies which attempt to understand the effectiveness of various interaction lockout mechanisms. In particular, we are interested in how a varying intensity level of an ILM is adhered to by users, and how users strategically leverage an ILM as a commitment device to achieve their goal of self-regulating smartphone use. Furthermore, it is important to understand user experience with ILMs that may negatively influence the willingness and persistent use of the intervention app.

We developed GoalKeeper, a smartphone app that applies the concept of a commitment contract that binds users with their voluntary self-defined daily goals of a usage time limit. If users exceed the self-defined time limit, an ILM intervenes in their smartphone use. As a comparative experimental study, we developed three versions of GoalKeeper with different levels of ILM intensities:

- *Non-ILM* delivers persuasive notification dialogs which constantly remind users of their time limit goal after they exceed their time limit goal.
- *Weak-ILM* locks users' phones for an increasing amount of time (e.g., 1, 5, 15, 30, and 60 min) after they exceed their time limit goal, but it still allows the use for a certain time period after lockout (e.g., a 15-minute allowance time after a 5-minute lockout).

- *Strong-ILM* locks users' phones after they exceed their time limit goal, and it does not allow the use for the rest of the day (until midnight).

We conducted a four-week-long field experiment with a within-group study design with the three aforementioned versions ($n = 36$) to answer the following research questions:

- (RQ1) How do people set goals with varying intensities of ILMs?
- (RQ2) How effective is each ILM in reducing smartphone use time and supporting the daily work focus of the user?
- (RQ3) What is the user experience with ILMs and how do people adapt to each of them?

Our main findings showed that strong-ILM can effectively reduce smartphone use most significantly. Not surprisingly, non-ILM was found to be least effective, as our participants often ignored the notification dialogs informing them of their excessive use and continued to use their phones. When our participants use the weak/strong-ILMs approaches, we found that they were more mindful about their goals and current states and attempted to use the remaining time allowance strategically. Our comparative experiment revealed that the participants' preferences were in the order of weak-ILM, non-ILM, and strong-ILM, demonstrating that users considered a certain level of restraint to be better than non-restrictive approaches of simply "informing" users. Moreover, flexibility when applying ILMs according to individual preferences as well as the usage context was found to be important with regard to continuance intentions. Our work provides valuable insights into the usefulness of supporting restrictive self-regulation. We recommend that researchers and practitioners to further explore design spaces of behavioral restriction in commitment devices to enrich intervention techniques in the persuasive technology domain.

2 BACKGROUND AND RELATED WORK

2.1 Commitment Devices and Temptation Management

Many people are interested in improving their problematic behavior. However, many fail to follow through with their initial goals [50]. This occurs mainly due to the inconsistency between the two selves, where one prefers instant gratification and the other values future utility more [21]. According to behavioral change theories, commitment devices, also sometimes referred to commitment contracts, are widely used to overcome this inconsistency between these two selves by enforcing voluntary restrictions on oneself until the goal accomplishment or binding penalties with the failures of accomplishing their goals [23, 63]. Some examples of commitment devices are binding a monetary loss with gym attendance [64], or leveraging one's social reputation by declaring a commitment in public [51]. There are commercial online services that help people to make such financial commitment contracts. Examples include stickK [69] and Beeminder [9].

These types of commitment devices can be categorized into hard and soft commitments [14]. Hard commitments involve actual tangible penalties for failure (e.g., a financial loss) while soft commitments have primarily psychological consequences. Alternative categorization is possible according to the mutability of the consequences [63]. Immutable commitment devices cannot be reversed by future choices, while mutable commitment devices can be reversed. An example of an immutable commitment device is a savings account on which interest is forfeited if a monthly deposit is not made, and an example of a mutable commitment device is the purchasing of food in small portions to reduce food intake. However, the consequences of the latter can be easily modified by going back to the grocery store and buying more food.

The theoretical basis of commitment devices is grounded in behavioral economics [23, 63]. However, little is known about how such commitment devices are used for technology-based interventions and their impacts on behavior changes, particularly in the context of interaction lockout mechanisms. Such technology mediated commitment devices differ from traditional commitment devices in that interactive systems may serve as *social actors* (as noted by B.J. Fogg [20]) that adaptively react to user interactions and impose certain penalties or restrictions on the user (e.g., locking out smartphones after limits on use are exceeded). In other words, users self-configure the commitment devices in an interactive system and delegate the enforcement to the interactive system which constantly tracks user behaviors and levies the associated penalties if the users violate the contracts. Given that the executing party of commitment devices is external (i.e., interactive systems), interestingly, users may feel coerced when the penalties such as lockouts are imposed by the system.

2.2 Persuasive Designs for Supporting the Smartphone Overuse Problem

In recent years, there have been many studies about intervening in smartphone overuse. Increasing self-awareness of usage behavior is one of the most popular intervention features. Many earlier studies provided self-monitoring by summarizing and visualizing a user's smartphone usage (e.g., usage time per day, frequently used apps) [40]. MyTime [26] more actively helped users become more aware of their usage via repeated notifications. It allows users to set daily usage goals for specific mobile apps and intervenes in cases of overuse by consistently sending timeout messages if usage goals are violated. Likewise, Good Vibrations [54] nudges users to regulate overuse behaviors with a *subtle*, repeating phone vibration.

Some studies harness social supports for smartphone intervention. NUGU [38] allows a group of users to share their usage and limiting behaviors. Researchers have found that social support can be helpful when attempting to maintain anticipated behaviors by helping users continuously to self-judge and allowing them to compare their usage times socially [38]. Lock n' LoL [37] also creates positive tension among group members in a place by explicitly sharing the members' current usage times in realtime. FamiLync [36] supports parent-child interactions for parental mediation on smartphone use by leveraging a public space for sharing usage information. Plan & Play [27] allows parents and children to collaboratively plan their use of apps (e.g., game choices and their playtime), showing that goal setting helps facilitate self-regulation. Home Trivia [19] is a home activity tracking system of digital technology use with gamification for better support of self-reflection among family members. Dong *et al.* [19] found that Home Trivia fostered social interaction and induced a positive use of technology. Interestingly, Ghosh *et al.* [22] reported that there is considerable tension between parents and children (i.e., safety vs. surveillance) with existing intervention apps.

Smartphone notifications are known to affect smartphone use times and usage habits. A study by Oulasverta *et al.* [56] showed that smartphone checking behavior comprises a large part of smartphone use time and that habits may increase overall use. Dabbish *et al.* [18] found that such external interruptions are likely to be followed by self-interruptions which may become habituated. A Ubicomp community approached this problem with an interruptibility management systems to detect the opportune moment for notification delivery [58]. Abdullah and colleagues found that there is a cognitive rhythm of alertness influenced by multiple factors including chronotype, sleep pressure, and social obligations [2]. They suggested that the unobtrusive biological sensing of the cognitive rhythm can be applied to the design of a tool to prevent cyberloafing or to encourage the productive use of technology. A study by Pielot *et al.* [59] found that disabling all notifications for a day decreased engagement with smartphones and increased productivity, but also observed some anxiety and more frequent checking behaviors. These works suggest the potential of technology and serve as a design guideline for successful intervention schemes for user's negative smartphone usage behavior.

Finally, several studies have examined various intervention systems that support the voluntarily use of smartphone blocking [36–38]. Some apps allow users to set rule-based or contextualized triggers for personalized blocking such as rule-setting [43] and location-based reminders [30]. While self-monitoring and social support features are designed to motivate users to change their behaviors, this type of usage blocking is more direct and explicit as it forces users not to use their smartphones, despite that fact that they are voluntarily set.

Blocking usage is a prevalent feature in smartphone intervention. However, to the best of our knowledge, none of the prior studies examined the effects of the lockout intensity level in relation to smartphone intervention or the design guidelines of interaction lockout mechanisms. For example, there is a lack of knowledge about how different levels of lockout intensity affect users' usage/limiting behaviors and how users develop strategies to deal with the lockouts. This paper tackles this research problem by exploring design spaces for lockout intensity in the area of smartphone intervention.

2.3 Restrictive and Coercive Interventions

The HCI and Ubicomp community expended an enormous amount of effort to design persuasive behavior change technologies. Nonetheless, knowledge about how to use restrictive and coercive methods for building behavior change technologies is lacking. One possible explanation is that such interventions may not be coherent with traditional behavioral change theories such as Self-determination Theory (SDT) [65] and the Theory of Planned Behavior (TPB) [5]. These theories strongly emphasize the importance of autonomy and control of one's own behavior in predicting one's behavioral intentions. Meanwhile restrictive and coercive interventions take just the opposite stance by reducing or eliminating one's autonomy. These approaches are meaningful when they provide a greater benefit (to oneself or for the public) in exchange for reducing one's autonomy.

Tromp *et al.* [70] proposed a two-dimensional influence model of behavioral changes in product design by considering the force of exerting behavioral influence (e.g., strong vs. weak) and the salience of influence attempts (e.g., implicit vs. explicit). According to this model, both persuasive and coercive designs make explicit attempts for behavioral changes, but a coercive design exerts much higher force than a persuasive design. As an example of coercive design, traffic spikes create a barrier to enforce a directional moving flow such as the entrance or exit to a parking lot. Drivers are discouraged from moving against the directional flow by perceiving the existence of the traffic spikes. Accordingly, the coercive design *externally* motivates people and forces them to encourage or discourage certain behaviors. In this work, we extend this notion of coercive design by demonstrating a self-imposed commitment device can be considered as a coercive design. As illustrated earlier, when signing a commitment contract with an interactive system, a user partly delegates the control authority of user interfaces to the interactive system, which then serves as a social actor that forces the user to maintain a positive behavior (e.g., a self-set goal of smartphone use limit). If a system's punishment (e.g., interaction restriction) cannot be avoided (i.e., immutable consequences such as screen lockout), it is likely that the user is *externally* motivated as well. Overall, commitment devices generally embody coercive design elements depending on how they are configured (e.g., when and how to enforce interaction restriction) by whom (e.g., oneself or external entities). After a user opts in, signed commitment contracts can be generally considered as coercive designs as well.

Recently researchers have started working on building and theorizing behavioral change technologies based on restrictive and coercive interventions. First, Benford *et al.* argued that as long as they are carefully and ethically engineered *uncomfortable interactions* (involving emotional or physical discomfort) can be an important design tool that can help achieve positive-long term goals related to various application domains such as entertainment, enlightenment, and sociality [10]. Uncomfortable interactions can be used to realize behavioral changes. Rekimoto *et al.* conceptualized inconvenient interactions by which an interactive system encourages or requires a user to perform some actions as long as the actions may bring about long-term benefits to the users (e.g., healthy eating, weight management) [61]. For example, the *Inconvenient Microwave* approach works only if a user engages in some exercise with a step machine connected to the microwave. This approach facilitates temptation bundling in that users are only allowed to have gratifying experiences only on occasions during which they show goal-consistent behaviors (e.g., exercising) [49]. Inconvenience interactions are also applicable to social groups. For example, Shin *et al.* designed a posture change system based on social support and discomfort by having a user's bad posture cause mild physical discomfort to the helper (e.g., shaking the phone to unlock) [68]. Cox *et al.* illustrated that decreasing fluid user interaction with *design friction* can increase mindfulness during user interactions [16]. They introduced an intervention referred to as a *microboundary* that places a small obstacle prior to an interaction to prevent mindless user interaction, such as by introducing short lockouts in number input tasks to minimize input errors. These prior studies contribute to the introduction and conceptualization of restrictive and coercive interventions. However, they did not explore how to design restrictive interventions in the contexts of commitment devices (e.g., goal setting vs. restriction intensity).

There have also been applications that employ restrictive interventions targeting smartphone and media usage reduction. PomodoLock uses a timeboxing technique with a app/website blocker to restrict access to user-defined counterproductive sources across multiple devices to minimize self-interruptions in the workplace [31]. Coco's Videos uses a *lockout* mechanism to control media consumption by children [25]. Park *et al.* bundled a cognitive task as a small barrier or microboundary to mitigate unconscious and frequent interaction with mobile devices [57]. Creaser *et al.* used phone blocking application which was activated when mounted in a car [17]. These studies have shown both positive and negative results from user evaluations. However, none comparatively evaluated how different types of lockout designs influence the usage behavior of mobile devices. We aim to fill this gap by designing various lockout mechanisms differing in terms of their lockout intensities and by providing a comparative understanding in this domain of research.

3 DESIGN

To understand the effectiveness of and experience with commitment devices with varying levels of lockout intensity, we designed and implemented the Android app "GoalKeeper." We extensively reviewed prior studies and commercial apps related to mitigating smartphone overuse to design lockout variations and to determine the common behavioral change components. We refined the design parameters of lockout mechanisms by interviewing nine participants (five female, mean age = 22.7, *SD* = 2.4). We showed them mockups of our initial mechanism design with the functional descriptions. This parameter-tuning process with the participants was particularly important to meet a reasonable level of usability in their daily use. The detailed app design rationale and features are described in the following sections.

3.1 Lockout Mechanisms of Commitment Devices

The core design components of a commitment device are 1) goal setting and 2) intervention mechanisms, or binding consequences upon a failure to achieve the goal. After thoroughly reviewing intervention mechanisms, we selected two types of interaction lockout mechanisms (ILMs) with varying intensities: denoted here as weak-ILM, and strong-ILM. For baseline comparisons, we also designed what is termed non-ILM, which simply warns a user without imposing any usage restrictions.

3.1.1 Goal Setting. According to prior studies, intervention apps typically ask users to specify a temporary lockout time [31, 37, 38], or to define the daily allowance time [26]. We consider setting daily goals for a week, which is appropriate for observing the effect of lockout mechanisms as a commitment device, as the willpower of self-control degrades over time [8]. Therefore, users are asked to define the maximum smartphone use time using the goal-setting feature in GoalKeeper (see Figure 1). There are two separate goal times for weekdays and weekends. This is intended to support the different daily activity patterns closely related to smartphone use. Our interviews with participants showed that they wanted to use their smartphones less on weekdays than on weekends, when they take a rest with less intention to reduce their smartphone use. Given that it is difficult to set daily goals due to lack of usage awareness, we collect week-long data as a baseline and suggest a personalized guideline during goal setting: i.e., a 10% or 20% reduction during weekdays and weekends.

Users set the goal on Mondays upon the start of the week, specifically a daily use limit for weekdays and weekends. After the users set the goal time, they are only allowed to modify the goal once during the week. This one-time-change rule was an important parameter derived from interviews, as allowing multiple modifications would nullify the intended effect of the lockout-based commitment devices. In contrast, some of our participants

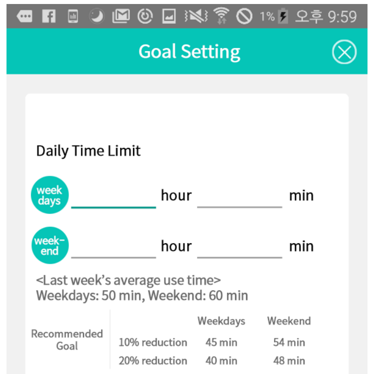


Fig. 1. Goal setting screen (goal recommendation based on baseline usage)

commented that they would not use the app if this type of goal modification is not permitted at all. Our interview results showed that one-time modification per week was deemed acceptable among the participants.

3.1.2 Non-ILM. Non-ILM was designed to remind the user of their self-defined daily limit without any lockout mechanisms. When the user hits the time limit goal, a dialog pops up which reads 'You have exceeded your time limit goal! Use time: x min, Time limit goal: y min, Will you continue using?' (Figure 2). This warning is similar to the guideline used in MyTime [26]. Our participants considered the notification interval to be an important parameter that influences their behavior. If the interval is too frequent, it would be too disturbing; if it was too sporadic, it would be ignorable. Our participants preferred a notification interval of 12.1 min on average (SD = 8.1). Based on the results of interviews, after the time limit goal has been exceeded, the notification dialog appears when a user unlocks the phone, or after every additional 15 minutes of use.

It is known that goal setting itself has a strong effect of goal achievement [44]. Moreover, previous studies of self-monitoring through interaction log visualization [36–38] and reminders with notifications [26] showed positive effects on behavioral changes. Though mechanism does not impose any physical interaction lockout on the user, it may provide a means of self-monitoring and warning mechanisms so that the users can quit using their smartphones.

3.1.3 Weak-ILM. Weak-ILM aimed to restrain the user's smartphone use via a light approach, specifically by a repetitive lock-unlock mechanism (Figure 3(a)). Once the user exceeds the time limit goal for the first time, GoalKeeper locks the phone for 1 minute, followed by a 15 minute allowance time. When the user spends the 15 minute allowance time on using the phone, the lockout will be imposed for the second time. In this case, the lockout duration will increase from 1 minute to 5 minutes. Again, after this lockout duration ends, another 15 minute allowance time is given. In our design, the lockout duration increments according to a predetermined schedule of 1, 5, 15, 30, and 60 minutes where 60 minutes is the maximum lockout duration. This pattern is somewhat similar to Android or iPhone's screen lockout protocols [55], which were designed to prevent repeated password input errors for security reasons.

We carefully considered two important parameters in Weak-ILM design: i.e., the lockout duration and the use duration (or allowance time) between subsequent lockouts. The allowance time was intended to provide *temporary use* after each lockout. The notion of *lightness* was translated into these parameters.

The suggested lockout duration from the interviewees varied from a minimum of 20 seconds to a maximum of 60 minutes ($M = 21 min.$, $SD = 18.6$). We referred to the screen lockout protocol of iPhone as well. Our participants mentioned that the lockout duration must be long enough such that they can move onto another

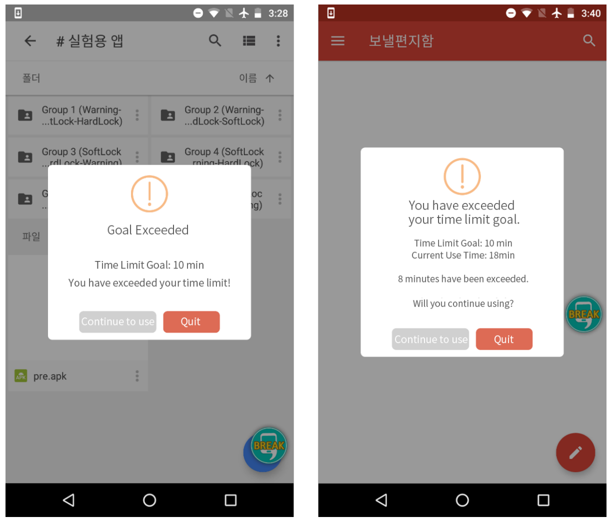


Fig. 2. Notification dialog of non-ILM, when the time limit goal was exceeded during use (left) or screen unlocking happened after exceeding goals (right)

task. Moreover, five out of nine interviewees positively agreed that this type of increasing lockout duration as on the iPhone would be helpful for self-regulating their smartphone use. Regarding the allowance time between subsequent lockouts, our participants considered that an appropriate allowance time ranged from 10 minutes to 60 minutes ($M = 19.1min.$, $SD = 19.3$). In our design, the allowance time was set as 15 minutes. This duration is consistent with that of non-ILM, which uses a 15-minute interval between each notification.

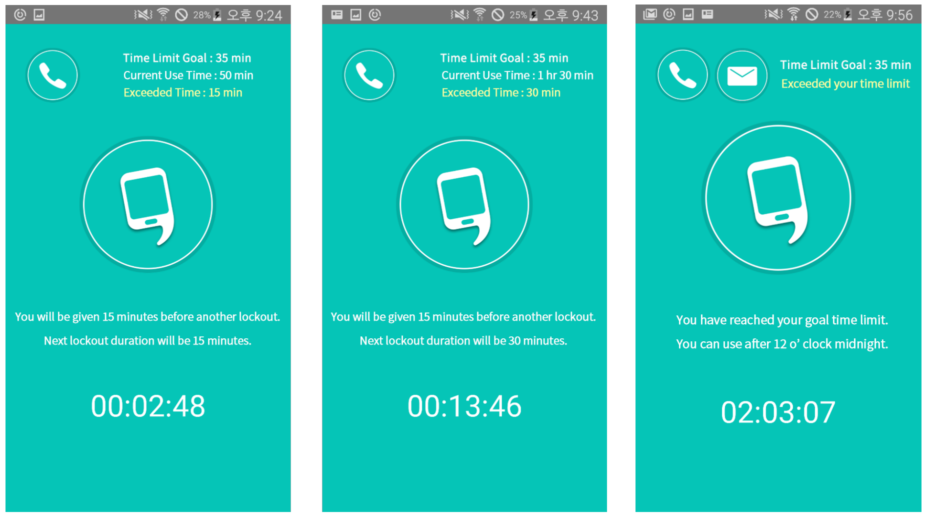
During the lockout period, we only allowed access to dialer and contacts apps, as they were considered an essential function of phones, and considering that emergency situation could arise (e.g., calling 911 or contacting family members or friends in case of an emergency). We concluded that other apps were less critical in that restricting those apps could not seriously affect their daily activities of users. Any other interactions including the physical home button were disabled.

3.1.4 Strong-ILM. Strong-ILM leverages an immutable lockout mechanism, which disables any smartphone use (except the phone calls or contact access) after the user hits the time limit goal (Figure 3(b)). The lockout persists until midnight at which the daily time count is reset. Our interview results showed that the participants may prefer to use strong-ILM if they are willing enough to reduce their smartphone use. Some participants noted that it may be too restrictive actually to use in their everyday lives. As shown later, we found that usage context and user needs mattered, and nuanced responses were observed in our field deployment.

3.2 Common Features

We embedded several basic behavioral change mechanisms as common features in all versions. These features include a timeline and statistics visualization, a usage notification bar for direct access to the goal and usage statistics, and a break mode for temporarily limiting smartphone use. These mechanisms were selected because they were commonly used in existing services.

3.2.1 Timeline and Statistics. GoalKeeper logs the user's application usage statistics in the background. The usage statistics are visualized with the timeline to enable a quick review of the user's smartphone usage history



(a) Lockout screen of weak-ILM (b) Lockout screen of strong-ILM

Fig. 3. Lockout screen after exceeding the time limit goal

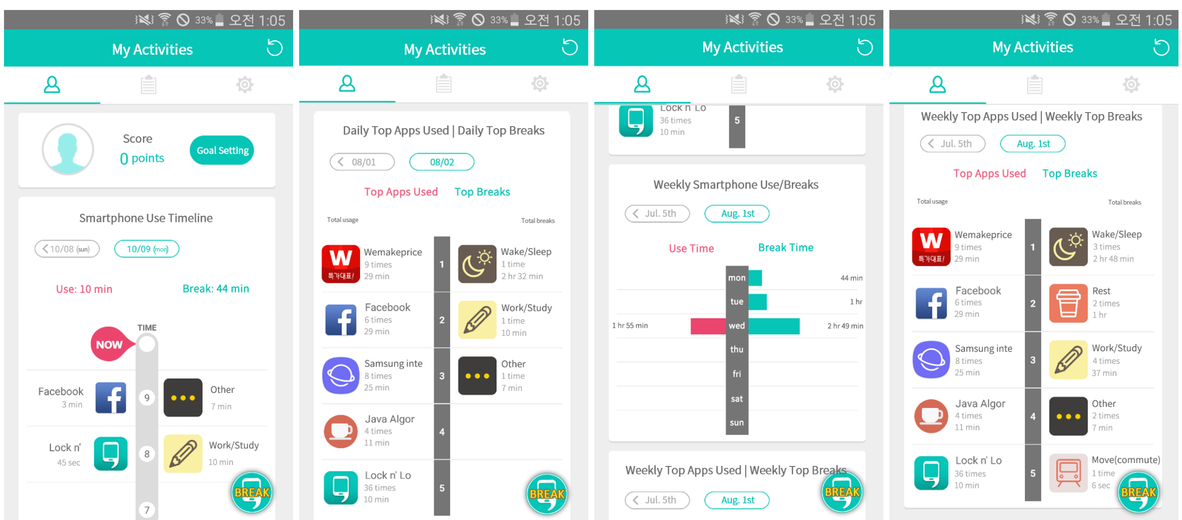


Fig. 4. Timeline and Usage statistics screen

(Figure 4). This type of timeline-based visualization scheme was also employed in previous studies [30, 36, 37]. The timeline provides an opportunity to self-reflect, which may lead to self-regulation. Further, time statistics by day as well as the week's top five apps are displayed. These features all contribute to amplifying the user's understanding of their smartphone usage patterns.

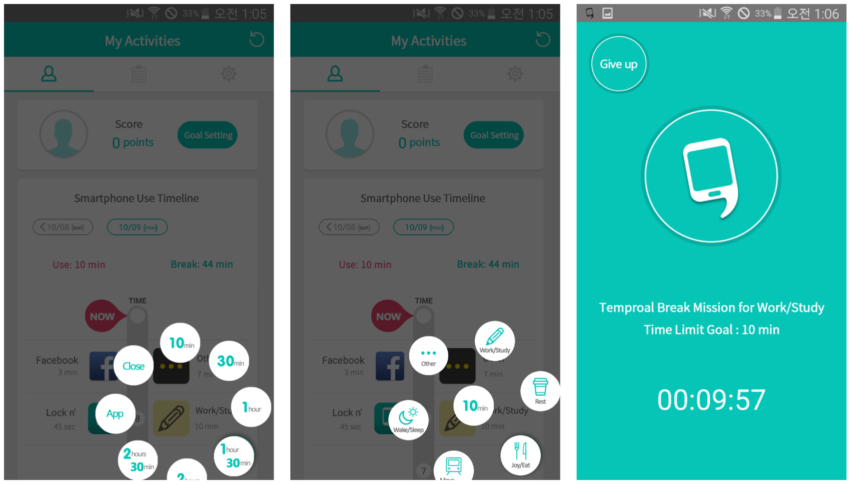


Fig. 5. Temporary break mode: selecting time (left), task (middle) and temporary lock screen (right)

3.2.2 Usage Notification Bar & Active Notification. The Android OS provides a notification bar that can be accessed without entering the application. We provided users' time limit goals and their total use times on the notification bar so that users could have quick and easy access to their usage information. This is a passive means of informing them of their current status. We also included an active notification bar that informs the user 30 min and 5 min before they hit the time limit goal.

3.2.3 Break Mode. The break mode is a timer-based, temporary self-limiting commitment device (Figure 5). It differs from the weak- and strong-lockout mechanisms in that it does not impose any physical lockout on the use of the smartphone, and the user must self-trigger the mode whenever it is needed. The user clicks the break mode button to choose the self-limit time and goal task (e.g., locking the phone for thirty minutes for studying). The application enters the lockout screen, displaying the remaining time for the self-limit. The user can exit the break mode at any time. This feature was documented as useful for self-limiting in prior studies [31, 37, 38].

We embedded this as a common feature to understand how useful it could be to use a self-triggered, voluntary break mode, along with proactive, system-enforced lockout-based commitment devices. As alluded to in a prior study [31], we hypothesize that such a self-triggered, voluntary approach would not be actively used by the users compared to proactive, system-enforced lockout-based commitment devices due to forgetfulness and a lack of motivation over time.

4 EXPERIMENT DESIGN

We conducted a four-week field deployment study of GoalKeeper with 44 participants to understand how different lockout intensity mechanisms of commitment devices impacted 1) goal setting, 2) smartphone use and daily activities, and 3) acceptance of such restrictive commitment devices.

4.1 Participants

Participants were recruited from a large university community portal. The inclusion criteria were Android phone users who considered themselves as excessive smartphone users and who were willing to reduce their use. We surveyed applicants' Transtheoretical Model (TTM) of behavior change stage [60] which has been widely used for assessing an individual's readiness to act on a new healthier behavior. This was to restrict study participation to those whose TTM stage was either two (considering usage reduction) or three (ready to reduce smartphone use). Thirty-two out of 46 initial applicants were at stage two, agreeing to the statement "I am worried about my smartphone usage time, but I don't have any detailed plan to control it." Twelve applicants were found to be at stage three, reporting "I have made some attempts to control my smartphone use, and am willing to take actions in the near future." Two applicants at stage one, reported "I do not feel the need to control my smartphone use time" and were excluded from participation. A total of 44 participants (mean age = 21.7, SD = 3.4) were finally selected. Each of them was compensated with approximately 45 USD for their participation.

After four weeks of the experiment, we were able to use the smartphone usage logs from the 36 participants. From the initial 44 participants, four participants' baseline usage logs were missing due to an application error. Another four participants' data were discarded because they found a workaround for the lockout mechanism, resulting in the abnormal use of GoalKeeper.

4.2 Procedure

We conducted a within-group study to understand the different effects of each commitment device comparatively. The experiment lasted four weeks, with the first week serving as a baseline measurement for regular smartphone use time logging. In the second, third and fourth weeks, participants were given different versions of GoalKeeper (non-ILM, weak-ILM, and strong-ILM) each week. In order to minimize the carryover effect [7], we used a complete counterbalancing design for the three conditions, assigning participants into one of the six groups with a different version-ordering.

4.2.1 Orientation Meeting and Pre-survey. Participants were given full descriptions of the GoalKeeper application and were instructed to install the application on day 1 of the experiment. They were also asked to complete smartphone use behavior questionnaires consisting their week and weekend smartphone usage times, the desired reduction ratio, and the smartphone addiction scale [39].

4.2.2 Weekly Use and Assessment Survey. Participants' smartphone activities were logged during the first baseline week, without any interventions. Based on this data, we recommended the two time-limit goals of 10% and 20% reductions from the baseline usage (Figure 1). From the second week, randomly assigned groups were given different versions of the GoalKeeper application in a counterbalanced order. At the end of each week, they were given a survey to record their user experiences with the version that they used. The main questionnaire also asked whether it was useful for mitigating the users' smartphone use, the levels of frustration/coercion they perceived, how focused they were (Focused Immersion Scale [4]), and how much of a workload they perceived (NASA-TLX [24]).

4.2.3 Exit Interview and Survey. They were surveyed on self-efficacy, self-control, the smartphone addiction scale, and their application version preference. The semi-structured interview was conducted to explore how each version supported the user in their efforts to achieve their goal, whether they had any negative emotional experiences, and whether habitual behavior on smartphone use changed. After each interview, we transcribed the data and conducted a thematic analysis [13].

5 RESULTS

We analyzed the quantitative usage log data as well as the qualitative data from the pre/post interviews. In addition, the weekly assessment surveys were analyzed to answer our research questions.

5.1 Goal Setting Patterns with Varying Lockout Intensities (RQ1)

Summary of the results

- More people changed their goals with a stronger lockout intensity of commitment devices.

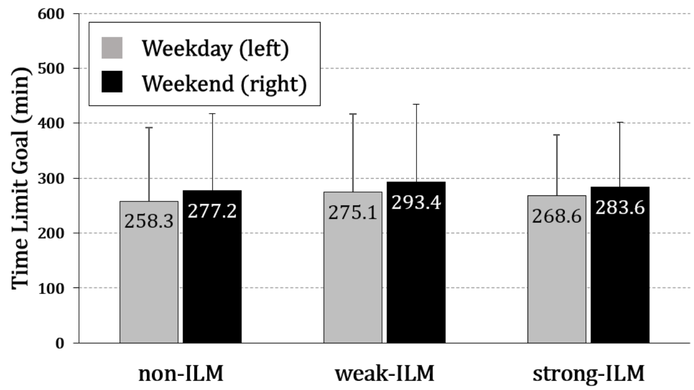


Fig. 6. Time limit goal with each type

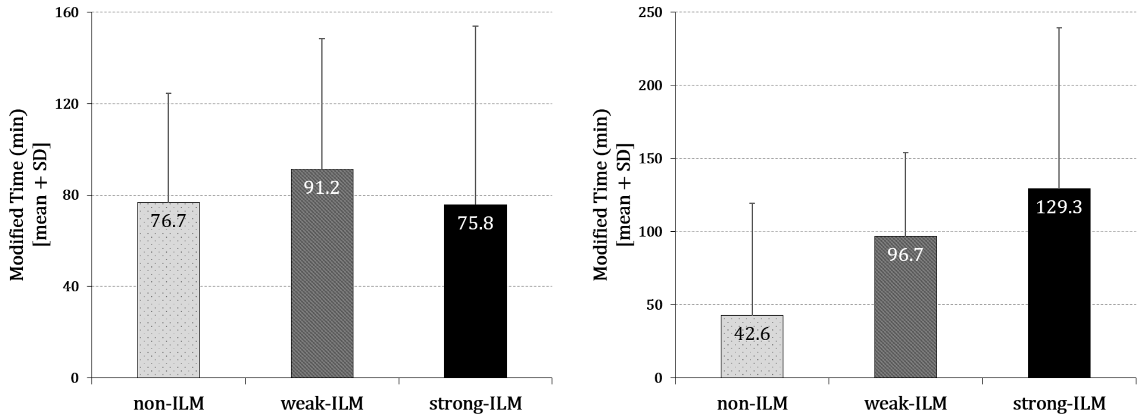
- Goal changes were mostly triggered when users encountered out-of-routine events or engaged in entertainment apps.

5.1.1 Time Limit Goal. Initially, we wanted to check how time limit goals were set in each condition. The results of the final time limit goals with each condition are shown in Figure 6. Repeated-measures ANOVA results showed there is no statistical difference among them for weekdays [$F(2,70) = 2.537$, $p = .094$, $\eta 2 = .130$] or weekends [$F(2,70) = 1.331$, $p = .278$, $\eta 2 = .073$].

In the exit interviews, we asked the users about the key criteria for setting their time limit goals. Our thematic analysis revealed that a user's goal-setting decision was primarily affected by the lockout presence, or by the personalized time limit recommendation. Although we did not find statistically significant differences in time limit goals across the different lockout intensities, our thematic analysis showed that the lockout presence affected 50% (n=18) of the participants' goal-setting behaviors. We further analyzed the time limit goals of those individuals who reported they considered the lockout presence as the primary criterion for goal setting. When considering this group only, we found that the repeated measures ANOVA results showed a statistically significant difference in time limit goals ($F(2,34) = 4.302$, $p = .022$, $\eta 2 = .202$) with a large effect size. Post-hoc comparisons with Bonferroni correction indicated statistical differences between the non/strong-ILMs ($p = .034$), but not between non/weak-ILMs ($p = .146$) nor between weak/strong-ILM ($p > .999$). These results indicate that the overall statistical significance of goal differences was affected by the mixed groups of participants who used different goal setting strategies.

Goal setting decision-making considering the lockout presence: 50% (n=18) of the participants reported that the main criterion for setting the time limit goal was the presence of restrict lockouts. The fact that the non-lockout version allowed the continued use of the smartphone even after the time limit was exceeded led them to set *tighter goals. "The consequence of exceeding the time limit with the non-lockout version does not cause much trouble. So I intentionally set a tighter goal."(P9) "In case of non-lockout, I reduced the time by 30 more minutes when compared with the time limit goal of the weak-lockout." (P33)*

Interestingly, the concern over being locked out of the smartphone with both the weak and strong-lockout mechanisms caused several participants to set a loose goal. Particularly, they showed greater *anxiety* with the use of the strong-lockout scheme. *"I was quite worried that I might not be able to use my smartphone, so I set a more loose goal" (P31) "I added an hour to the strong-lockout version compared to the weak-lockout version, just in*



(a) Weekday (b) Weekend

Fig. 7. Average of time difference between the original and the modified time limit goal

case (I was locked out)." (P22) "Once I experienced lockout with the weak-lockout scheme, I had to increase the time limit goal greatly." (P40)

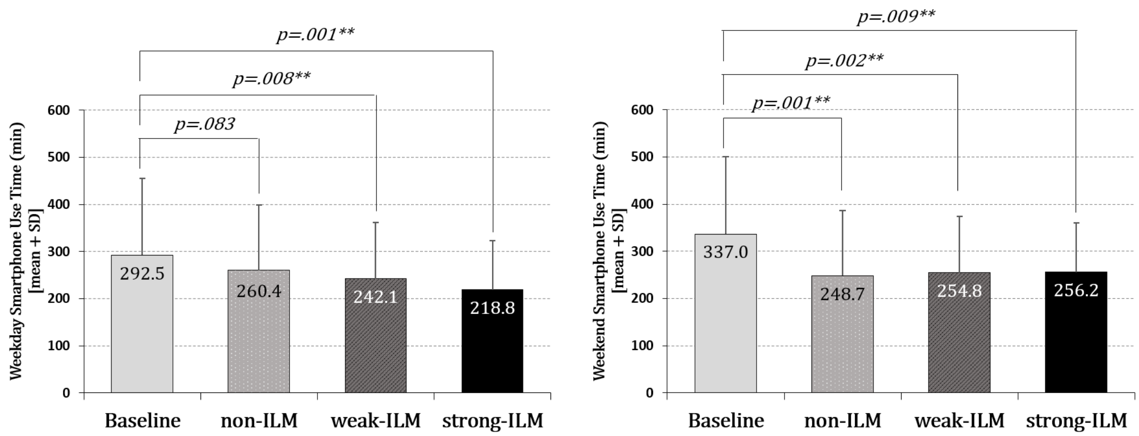
Goal-setting decision-making considering recommendation: 50% (n=18) of the participants either directly followed the recommended time limit goal, or customized their goals by referring to their usage levels and recommended goals. These participants were relatively unconscious of their regular usage behavior and showed less sensitivity to the lockout. *"The app would help me reduce my usage time, so I subtracted 30 minutes from the recommended amount." (P12) "I was still not able to grasp how much I used the phone the week before, so I just followed the recommended goal." (P3)*

5.1.2 Goal Time Modifications. In contrast, 15 participants did so with the weak-lockout version, while 20 participants with the strong-lockout version increased their time limit goals during the experiment. We conducted a closer examination of these time modifications from the original goal. Our results showed that weekday modifications did not differ greatly across the conditions (Figure 7(a)). However, the differences in the weekend goals, particularly between the non-ILM and strong-ILM conditions were large (Figure 7(b)).

These participants explained the reasons behind their modifications in the post-interview. First, these modifications mainly occurred when the participants were out of their routine, such as on a business trip or visiting friends in other cities. During such periods of time, they had to, or wanted to, use their phones more.

"When I was out for a long time, there were frequent situations in which I had to communicate with my peers, so I modified the goal." (P14) "When I was taking a bus to Seoul, and I would have to spend at least four hours both ways, I modified the goal beforehand." (P4) Some participants had to use a navigation app for driving. *"It was my first visit to this shopping mall, so I used a smartphone navigation app. My phone became locked on my way back, giving me a hard time." (P7)* These out-of-routine incidents caused some of our participants to increase their time limit goals.

In addition, extensive consumption of entertainment apps contributed to the modification behaviors. *"The game I was playing was quite addictive, made me play more and more. Later, I had to increase the time from what I initially planned."(P16) "I was watching some sports game on my phone. The temporary lockscreen (weak-ILM) appeared, so I extended my goal by three hours. (P27))*



(a) Weekday average use time

(b) Weekend average use time

Fig. 8. Use time against baseline values

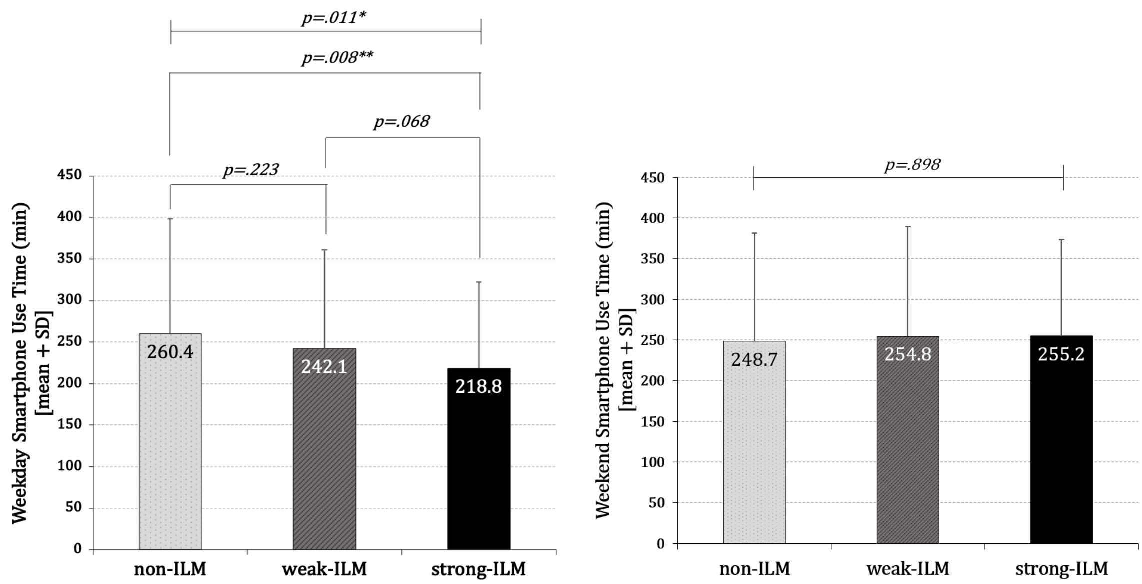
5.2 Effectiveness in Reducing the Use Time and Supporting Better Work Focus (RQ2)

Summary of the results

- Participants' smartphone use time decreased by a statistically significant amount in all conditions except for the non-ILM in the weekdays.
- Smartphone use time decreased as the lockout intensity of the commitment device increased, and this tendency was strengthened in participants who did not modify their goals.
- There were no statistically significant differences between each mechanism in terms of its support of the users' daily work focus.
- However, the participants considered the temporary break feature as a factor contributing more to their daily work focus over the lockout mechanisms.

5.2.1 Smartphone Use Time. Smartphone use time is the most direct way of measuring the effectiveness of lockout mechanisms. First of all, we examine whether each treatment successfully reduced the usage time relative to the individual's baseline usage using pair-wise comparisons. We then investigate the effects of lockout intensities by comparing three treatment conditions.

We initially compared participants' weekday baseline times with the time spent on their smartphone when each of the three lockout mechanisms was given. With the non-lockout version, participants reduced their time by 32.1 minutes per day. However, when we analyzed the outcomes with a paired t-test, the difference was not statistically significant ($t(35) = 1.739$, $p = .091$, $d = 0.21$). Meanwhile, the smartphone use times of the weak and strong-lockout versions decreased by 50.4 minutes ($t(35) = 2.801$, $p = .008$, $d = 0.35$) and 73.7 minutes ($t(35) = 3.507$, $p = .001$, $d = 0.54$) respectively. This implies that the stronger the lockout intensity is, the more the user can reduce their time. The weekend smartphone use time between the baseline and each experimental weekend also showed the following statistically significant decreases: non-ILM ($t(35) = 3.656$, $p = .001$, $d = 0.56$), weak-ILM ($t(35) = 3.440$, $p = .002$, $d = 0.51$), and strong-ILM ($t(35) = 2.755$, $p = .009$, $d = 0.53$) as shown in Figure 8(b).



(a) Weekday average use time

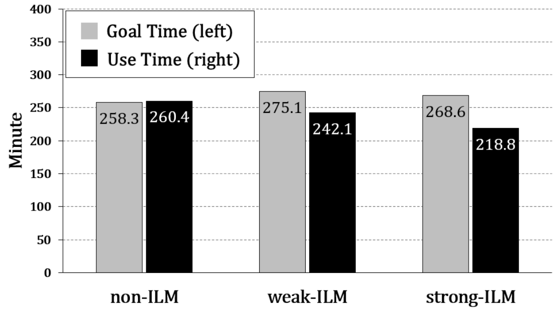
(b) Weekend average use time

Fig. 9. Smartphone use time comparisons among the three experimental conditions

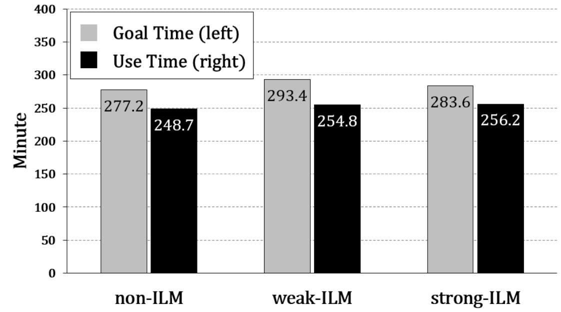
We then checked whether there were statistically significant differences in smartphone usage times over weekdays/weekends across the three conditions. The results of a repeated-measures ANOVA showed that there were statistical differences among the three conditions for weekdays [$F(2, 70) = 5.114$, $p = .011$, $\eta 2 = .231$] with large effect size. Post-hoc comparisons using Bonferroni tests indicated statistical differences between the non/strong-ILMs ($p = .008$, 95% CI [9.129, 74.038]), marginally significant differences between the weak/strong-ILMs ($p = .068$, 95% CI [-1.299, 47.949]) but none for the non/weak-ILMs ($p = .223$, 95% CI [-6.710, 43.226]). The results of a repeated-measures ANOVA with weekends ($F(2, 70) = 0.108$, $p = .898$, $\eta 2 = .006$) were not statistically significant.

5.2.2 Goal-Use Discrepancy Analysis. We analyzed the discrepancy between the self-defined goal time limit (the final goal, regardless of modified or non-modified goals) and actual smartphone use with the three versions of GoalKeeper. We observed that the discrepancy increases as the lockout intensity increases during weekdays (Figure 10(a)). For the non-lockout mechanism, the goal-use discrepancy for weekdays was -2.1 minutes, implying that the participants tended to set tight goals. In contrast, the goal-use discrepancy of the strong-lockout mechanism was relatively large (49.8 minutes). Meanwhile, weekends showed similar goal-use discrepancies across the conditions (Figure 10(b)). The goal-use discrepancies in this case were as follows: non/weak/strong-ILMs were 28.3, 38.6, and 27.4 minutes respectively. These results show that no conditions exceeded the time limit goals on average, except for the weekday non-ILM condition.

5.2.3 Perceived Daily Work Focus. While smartphones are often used as a productive tool, there are many daily activities in which smartphone usage could be distractive; e.g., working and studying, having conversations with friends and family, and even just taking a rest. We hypothesize that Goalkeer helps our participants to better manage distraction by self-regulating smartphone use.



(a) Weekday Goal time vs. Use time



(b) Weekend Goal time vs. Use Time

Fig. 10. Goal-Use discrepancy

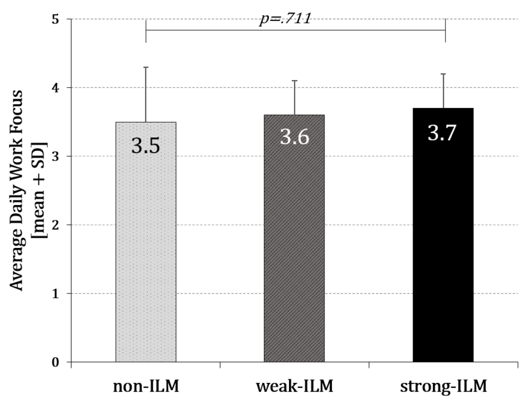


Fig. 11. Daily work focus

We measure and compare how different mechanisms embedded into commitment devices would support users' daily work focus. In the weekly survey after using one version of GoalKeeper, we asked whether the particular type of a lockout mechanism that the user experienced supported their efforts to focus on their daily work. We used the Focused Immersion questionnaire [4] for this measurement.

The non-lockout version scored 3.5/5.0 (SD = 0.8), the weak-lockout was 3.6/5.0 (SD = 0.5), and the strong-lockout scored 3.5/5.0 (SD = 0.5) (Figure 11). There were no statistically significant differences among these three versions according to a repeated-measures ANOVA test. The LSD post-hoc test showed no differences between each pair as well. However, the absolute value exceeding 3.0 means that our participants leaned toward the positive side.

Consistent in the exit interview, participants mentioned that the time limit and the lockout mechanisms did help to reduce their total smartphone use times. However, they did not always help them to focus on their work or study. This occurred because when they were highly immersed in smartphone use, it was difficult for them to stop using the smartphone or at times they even failed to realize their goals. *"I didn't come to think of the time limit goal I had to keep, so it didn't really help." (P10) "It did not occur to me that I had this app installed at that moment." (P14)*

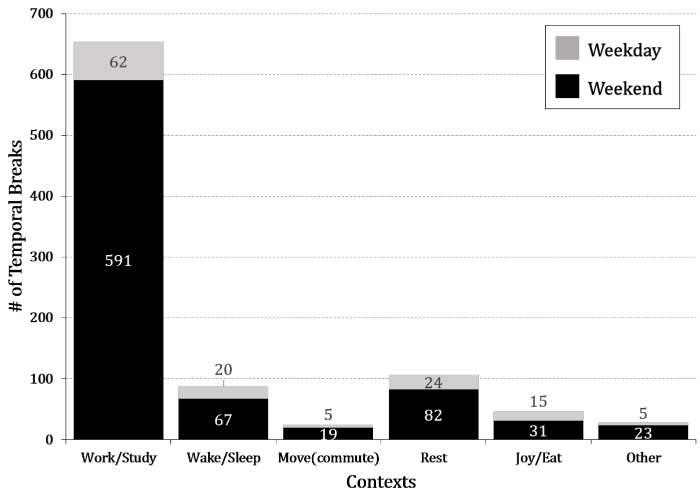


Fig. 12. The specific contexts in which the temporary break feature was used

Those who agreed that the goal limits were helpful, stated that this was the case only when using the weak or strong-lockout mechanisms together with the goal. The lockout mechanisms made them mindful of the fact that their smartphones would be locked. They attempted to use their devices less and sometimes maintained a physical distance between themselves and their smartphones if possible. *"It was helpful. I tried to keep my phone away from me." (P15) "When the smartphone was locked, it made me forget about using it and work on what I have initially planned" (P5)*

The majority of participants mentioned that they found the temporary break feature to be more supportive in terms of task concentration (Figure 12). The instant feedback from the unlocked screen reminded them of their temporary break from smartphone use. *"I said 'oh, I put the phone in break mode,' when I unconsciously unlocked the phone, and I could endure the temptation to use it." (P10)* In addition, the mindful use of this temporary break feature in a specific context on which to focus was considered helpful. *"I used a one-hour break when I was doing assignments, ..., it was especially helpful in class." (P13)*

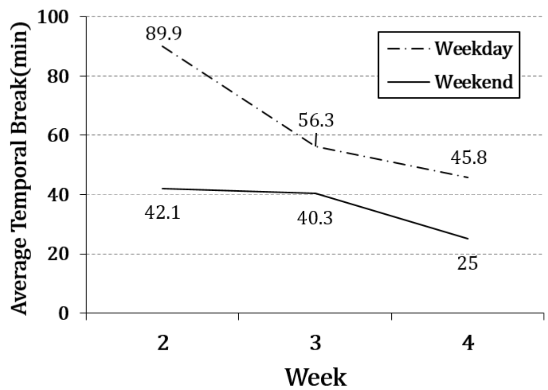
We analyzed the specific contexts with which the temporary break feature was used. Work or study was the dominant use case, followed by the resting, and then before sleeping. It is notable that 86.1% of the use of temporary breaks occurred on weekdays.

There were cases in which the break mode was said to be meaningless and even disturbing. *"Having to intentionally initiate the break mode was disturbing. I rather refrained away from touching the phone in the first place." (P14)* In addition, some participants well recognized its usefulness, but they either forgot to use it, or thought that it was cumbersome to use over time. The use of the temporary break feature on weekdays decreased from the first week ($M = 89.9min.$, $SD = 125$), to the third week ($M = 45.8min.$, $SD = 89.1$), as shown in Figure 13(a). The number of participants who used this feature also decreased (Figure 13(b)).

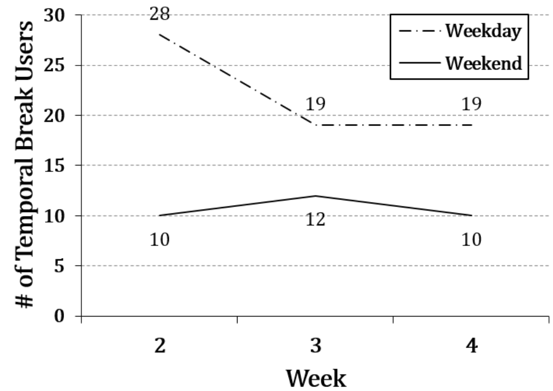
5.3 User Experiences and Preferences (RQ3)

Summary of the results

- The non-lockout version only gave a chance for self-reflection, and it was frequently ignored (92%).



(a) Average break duration



(b) Number of break mode users

Fig. 13. The use of temporary break feature

- The weak/strong-lockout versions introduced more inconvenience than the non-lockout version, but participants reported that these versions could effectively suppress the temptation to use their phones.
- Lockout conditions facilitated constant monitoring behavior of their smartphone usage time, as well as their planning behavior, but the strong-lockout version induced the greatest negative emotion.
- In terms of preferences, 52.8% preferred the weak-lockout version the most, the 33.3% the strong-lockout version, and 13.9% the non-lockout version, reporting that the simple warning was too easy to ignore and was less helpful for reducing their usage time.

5.3.1 User Experiences.

Non-lockout: The core mechanism of the non-lockout version was the repeated notification of a warning message when exceeding the time limit goal. This repeated notification gave users increased awareness of their goals. This then led them to engage in self-reflection, and experience regret and a sense of shame. *"I was unaware that I reached the time limit. It was then when I had cautiousness." (P16) "The notification gave me a sense of shame, which stopped me from using the phone." (P26)* Some participants appreciated the fact that the application gave them the autonomy of self-control, unlike the other two lockout mechanisms. *"I like how the non-lockout version just reminds me of my goal, so I can control my smartphone use." (P10) "[It] makes me think that I should use it less, without pressure." (P3)*

However, this notification approach had its downside. Our data logs showed that 92% of the participants pressed the "continue to use" button (instead of the "quit" button) when the notification dialog appeared. The majority of participants reported that the warning dialog was too easy to ignore and was less bothering, reporting that *"It's only a dialog message. It does not lock my phone or help me stop using." (P32)*, *"It was meaningless, I just kept using the phone." (P36)*, and *"I was then aware of how much I used it, but it didn't lock, so in many cases I just kept on using it" (P4)*

Weak-lockout: The temporary lockout mechanism amplified the participant's mindfulness to reduce their smartphone use. They considered this temporary lockout to be inconvenient, and wanted to avoid facing such situations. Therefore, when they approached the time limit goal, they refrained from using their smartphones, particularly in what they called *meaningless* YouTube and Facebook activities. When checking emails and communicating by instant messengers, they attempted to plan their use. *"Being locked out is so inconvenient. I had* *to be cautious not to be locked out again." (P35)* The temporality of the lockout mechanism made the participants "naturally" give up using their phones and move on to another non-smartphone task. *"I can't use my phone. I would naturally do something else." (P28) "If I start on my work, I continue for at least 30 minutes. So being locked out for one minute or five minutes, then I would not go back to my phone once it was locked, even after it became unlocked." (P32)*

Interestingly, some considered the "temporary allowance time" after the temporary lockout as a "reward" for not using their phones, instead of treating it as a "penalty." P8 commented, *"This was really helpful. But it would be even better if you gave me a longer allowance time as a reward if I wait longer."* The 15-minute allowance was appreciated, being the right duration for only doing what is actually important. *"It gives me a 15 minute chance to use it. During this time, I can do what I really need to do, or move the data (to another computing device such as a PC or a tablet)." (P9) "Used it for a really short time. Opening and closing the emails" (P1)* Overall, the perception of the repeated lockout and unlock process was nuanced in that it can be considered as a reward and a penalty at the same time.

A negative experience arose when being temporarily locked out, as it drew the users' attention and caused them to continue to check their phones. *"I kept on checking how much time was left until it would unlock." (P38)* Another comment was that the short lockouts were meaningless. *"I just waited until it unlocked and kept using it" (P37) "Waiting a short period did not bother me. I continued to use it without much thinking." (P21)* During the lockout, we found that some participants used substitute devices such as PCs and tablets, particularly when they felt that they really needed to do some extra work. *"I turned on my PC to use the PC version of the instant messenger ..., whatever I needed to do, I used the PC." (P4)*

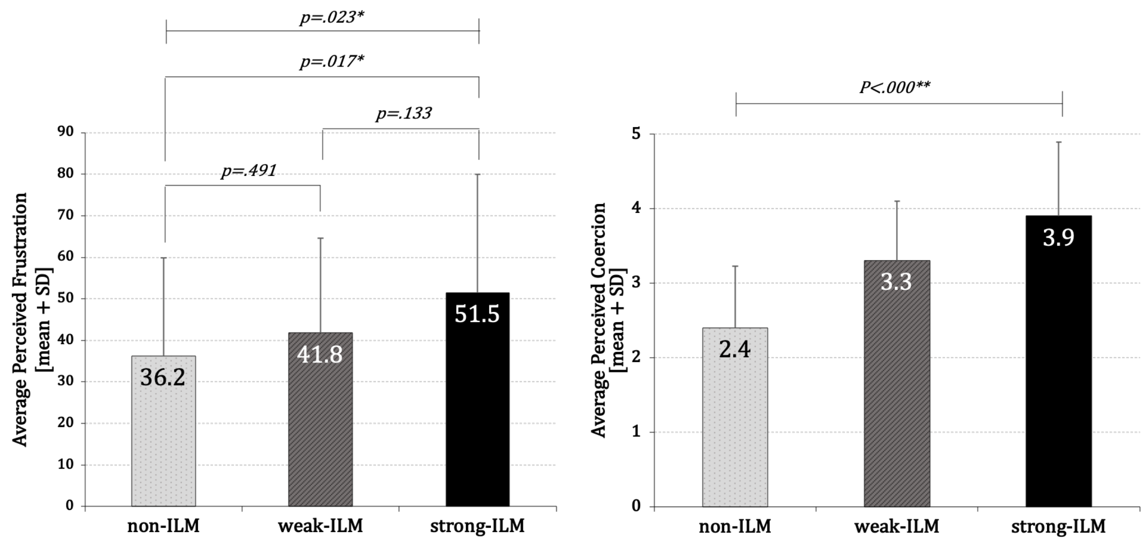
Strong-lockout: The strong-lockout version disabled all interactions with the smartphone until midnight if the time limit goal has passed. Participants perceived this as coercive and restrictive, but they agreed that it most powerfully reduced their smartphone usage time. *"Totally restricting use is effective" (P13) "The smartphone was locked, and I could not use it. The total use time naturally decreased." (P1) "It was hard, yet most effective in suppressing my temptation to use it." (P17) "It was helpful. I naturally gave up using it once I was locked out." (P28)*

Similar to the weak-lockout mechanism, the strong-lockout mechanism helped the participants carefully to plan not to exceed the time limit goal. Moreover, the strong-lockout version created a smartphone usage pattern of frequently checking the use time and planning the use accordingly. *"Once it is locked, I can't do anything with it. So, I had to break the time of use into small sessions, and manage the total time carefully." (P32) "I wanted to save some time for later use, when I really needed it." (P21)*

However, it was also perceived negatively in several aspects. First, the fact that it was locked presented a major burden. This burden was mainly expressed as displeasure. *"The total lockout was helpful, but I didn't feel good." (P35) "It was definitely useful for reducing the smartphone use time itself, but I don't think this is a good method." (P7)* In addition, there were some attempts to prevent the lockout. *"When the time limit goal approached, I tried to find a way to avoid the lockout." (P3)*

There were some situations in which the participants inevitably had to use their phones, but could not due to the lockout. In these cases, they usually expressed considerable negative emotions, such as discomfort, anger, and stress. Examples of such an inevitable situation occurred out of users' regular routines; e.g., using navigation apps when walking or driving, or receiving a certain request to handle a task. *"When I was going to another city by bus, I almost used up all of my goal time. Then, my friend sent me a map for the meet-up, which I could not check due to the lockout." (P10) "It was inconvenient. I had to write an experiment report, but the data was in the locked smartphone." (P9)*

5.3.2 Perceived Frustration and Coercion. One of the main concerns of the interaction lockout approach was that, it could introduce negative emotions. These negative emotions may cause the users to opt out or find a way to neutralize the system. Therefore, understanding how much of these negative emotions are perceived when



(a) Average perceived frustration (scale: 0-100)

(b) Average perceived coercion (scale: 1-5)

Fig. 14. Average perceived frustration and coercion

using the various levels of the lockout mechanisms is important. We assessed both perceived frustration and coercion with each version every week. These results are shown in Figure 14(a) and in Figure 14(b)

First, we measured *perceived frustration* using an item from the NASA-TLX assessment. We slightly modified the item as follows: "When you are using this application to achieve the goal of reducing your smartphone use, how much negative emotions, including anxiety, disappointment, pressure, and stress did you perceive?" On a scale of 0 (none) to 100 (very much), the non-lockout version scored 36.2 (SD = 23.7), weak-lockout scored 41.8 (SD = 22.8), and strong-lockout scored 51.5 (SD = 28.5). As the lockout intensity increased, the level of frustration increased as well. Our repeated-measures ANOVA showed statistically significant differences [$F(2, 70) = 2.227$, $p = .023$] with a small effect size of $\eta 2 = .199$. The Bonferroni post-hoc test showed a statistical difference between the non-ILM and strong-ILM ($p = .017$), but not between the non/weak-ILM ($p = .491$) nor the weak/strong-ILM ($p = .133$)

Next, out of the negative emotions, we particularly measured perceived coercion, which is a type of pressure imposed on a person by the system. It is a measure used to understand the likelihood of user compliance with a treatment or an intervention [66]. The results showed the non-lockout version was perceived as least coercive (2.4/5.0, $SD = 0.83$). As the lockout intensity of the lockout mechanisms increased, perceived coercion increased to 3.9/5.0 ($SD = 0.99$) for the strong-lockout version. Our repeated-measures ANOVA showed that the three versions were statistically different [$F(2, 70) = 20.129$, $p < .000$] with a small effect size of $\eta 2 = .542$.

In the exit interview, we asked the participants when these negative emotions were perceived and how it affected their behavior. The emotion before and after the time limit goal could be very different due to the actual experience of being restrained by the application. Thus, we explicitly asked about these two situational events.

Before exceeding the time limit goal: The non-lockout version did not place much pressure on the participants, as they were well aware that no physical lockouts would ensue. *"The notification dialog does not burden me* *much." (P12)* However, some participants make an effort not to exceed the limit. *"I did not think much, but I made a simple plan of how to use it ..." (P5)*

With the weak-lockout version, several participants were neither pressured nor stressed, reporting that they knew that the lockout would be released soon. *"The lockout can be released just after a while. It is not burdensome at all." (P15)* However, the majority felt some pressure and frustration due to the fact that they could be locked out. Thus, in order to avert a lockout, participants carefully planned their use. They frequently checked remaining time, and used PC or laptops for instead of their smartphone if necessary. *"When I checked the remaining time, I realized I really have to plan my use ..." (P1) "Consuming half of the time limit goal made me think I might not be able to read urgent messages. Therefore, I started saving up for later use." (P33)*

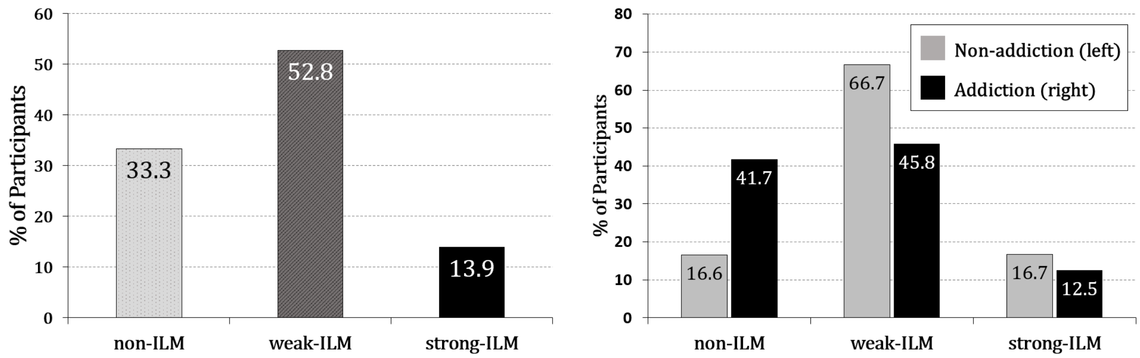
The strong-lockout version was perceived much more negatively relative to the other two mechanisms. A majority of participants expressed pressure, stress, anxiety, and sensitive emotions before reaching the time limit goal. Specific time of pressuring moments mentioned amounted to 10 to 20 minutes, and halfway to the time limit. Participants also stated that they exerted more effort not to be locked out compared to the other two versions. *"I planned to save 20 minutes for important calls or messages. That 20 minutes is really precious." (P15) "I tried not to use the phone much when passing halfway, as I was worried I would be locked out." (P33) "Really, it really locks, so I really tried my best not to hit the limit." (P38) "I had to worry about being locked out long before the limit. It's way too strong compared to the other two versions." (P30)*

After exceeding the time limit goal: Some of the reactions to exceeding the goal limit with the non-lockout version included feelings of disappointment, guilt and self-reflection that they could not keep their own promise. *"I felt disappointment and guilt in myself, and I stopped using my phone." (P26)* However, most of the participants reported that they naturally ignored the fact they exceeded their goal limit. Thus, our participants did not feel much pressure, and this did not lead to any reductions in their smartphone use. *"I pressed 'continue to use' whenever the dialog popped up. There was no pressure." (P20) "Upon the first notification dialog, I felt I should stop. Habit is something horrible. After several repeated uses, my burden and guilt disappeared." (P17)*

When using the weak-lockout version, participants expressed different levels of coercion and frustration when being locked out at an early vs. late hour of the day. This threshold hour was typically around 8 or 9 pm according to our participants. They reported that the fact that the reset time would arrive soon relieved their tension. *"Most of the lockouts were around 11 pm. Just one or two hours of waiting was fine." (P15)* However, when they were locked before this time, a greater burden and frustration were expressed. The lockout was temporal, but the fact that they could not use their phones at essential moments increased their stress level. *"Once it locks, I know it would be frustrating to be locked out again and again for longer periods." (P38) "It feels like I am living a time-limited life." (P1)*

The strong-lockout versions generated similar reactions in that the participants were greatly pressured when their phones were locked at an early time of the day. They both regretted the overuse to the point that they give up and accepted their restricted situation. The strong-lockout version was perceived very negatively because there were many "out of routine" incidents experienced by the participants, as noted in Section 5.3.1 describing user experiences. There were "give up and accept" behaviors, with one user reporting *"Oh, man ... I did use a lot today. I considered my phone missing from that time on." (P5) "Living like a primitive man. I just left my phone, giving up any attempt to use it for the rest of the day." (P33)* Regretting behaviors were also mentioned. *"Why did I use it late at night (after midnight, the usage time counts for the next day, leaving less to start that day with)." (P34)* A more positively accepted case was also mentioned. *"I felt rather good that I could not use my phone!" (P28)*

5.3.3 Preferences. A majority of the participants (52.8%) preferred the weak-ILM version (Figure 15(a)). These participants appeared to be satisfied with the balance between lockout and flexibility by comparing the weak-lockout version with the non-ILM version. One participant stated, *"it is well balanced between pressure and freedom." (P20)* Another participant mentioned, *"Its lockout intensity is the most reasonable among all methods.*



(a) Preference

(b) Preference differences addiction vs. non-addiction

Fig. 15. Average preferences of ILMs

... [with weak-ILM], I can plan and manage my smartphone use based on the expectation of possible outcomes or rewards ... though it gives some pressure."(P8)

Moreover, weak-ILM was preferred because it is more flexible than strong-ILM to deal with urgent usage contexts. One participant responded, *"Of course, it somewhat forces me not to use my smartphone. But, its level is not too strong, so I was rarely concerned about an emergency during which I would really need to use the smartphone." (P1)* Another participant mentioned, *"I understand that locking is a type of punishment due to my incorrect behaviors, so I willingly take it. However, it's also true that I'm very inconvenienced if I cannot use my smartphone completely." (P30)*

Next, a significant number of the participants (33.3%) reported that they most liked the non-ILM version. These participants mostly appreciated non-ILM because it helped them to be aware of their smartphone usage and to maintain their attention on their behaviors. One participant said, *"I want to keep using the weakest one (non-ILM) to check my usage behavior and to enhance my awareness of it." (P13).* Some of these participants also expressed strong concerns about potential inconvenience in the lockout contexts. Some participants responded, *"I think locking my smartphone will make it very inconvenient, so I prefer non-ILM" (P12)* and *"I prefer a weak pill at first and change it to stronger one depending on the situation after some time. ... Radical changes do not seem effective for me." (P12)*

In our supplementary analysis, we examined whether preference differences exist between the addiction and non-addiction groups (Figure 15(b)). Based on our addiction scale responses, we found that 24 participants could be categorized into the addiction group and 12 into the non-addiction group (based on the Smartphone Addiction Scale [39]). It is notable that non-ILM was most preferred by 41.7% of the addiction group (weak-ILM: 45.8%, strong-ILM: 12.5%), while only 16.7% of the non-addiction group preferred this version (weak-ILM: 66.7%, strong-ILM: 16.7%). This result indicates that people who heavily depend on their smartphone devices tend to feel more negative emotions caused by a strong level of lockout intensity.

Finally, strong-ILM was least preferred (13.9%), but many participants still mentioned its usefulness in specific contexts that require a high level of focus. One participant said, *"I need to use it (strong-ILM) during the exam period. It will surely be helpful because it guarantees that I will not exceed a certain usage time."* (P15) Another participant similarly reported, *"These days, I am not that busy. ... When I was rushed, the most helpful method was strong-ILM. With this method, I can certainly reduce the time I spend watching YouTube or Netflix." (P32)* However, most of the participants responded that they are less likely to use this method because it is too stressful. One participant mentioned, *"Surely, I think I need to reduce my usage time. Nevertheless, I don't want to be stressed for doing it." (P19)*

6 DISCUSSION

Managing the acceptance and effectiveness of ILMs presents two different matters that are closely related. An intervention application as a commitment device should be acceptable for users to use it persistently. It must also maintain a proper level of effectiveness in achieving its goal of reducing the user's smartphone usage time. A persuasive and flexible design can play a key role in balancing these two goals of ILMs. In this section, we discuss the key findings pertaining to the acceptance and effectiveness of ILMs and extract design implications for persuasive and flexible ILMs.

6.1 Persuasive Design for Effectiveness of ILMs

6.1.1 Motivating Users to Try More Challenging ILMs. Our results show that all three versions of GoalKeeper significantly reduced the smartphone usage times of the users (as opposed to the baseline). Furthermore, we found that the greater the ILM intensity, the greater the effectiveness of self-regulating smartphone use becomes, despite negative user experiences. Indeed, the strong-ILM version was least preferred by a majority of the participants. Ironically, those who were in the addiction group based on our survey scales largely (41.7%) preferred the weakest intervention method (i.e., non-ILM). In fact, this addiction group was likely to be more vulnerable to the pressure imposed by strong ILMs, but at the same time, this group is more likely to benefit from stronger interaction lockouts for behavioral changes. Therefore, it will be more helpful to track users' motives and ILM usage patterns, and then continuously guide them to adaptively vary the degree of lockout intensities, possibly challenging them by carefully considering behavioral change strategies (e.g., making small changes) [53].

6.1.2 Guiding to Better Goal Setting. In our main experiment, we logged the smartphone usage times of the participants during the baseline week. One pre-survey question asked the participants how much time they spent on their smartphones per day. We calculated the difference between their self-reported "expected usage time" from the survey and the "actual usage time" from the baseline usage log, finding that our participants underestimated their usage time, with average gaps of -35.9 minutes (SD = 165.8) for weekdays and -29.5 minutes (SD = 187.8) for weekends. To reduce this perceptual gap, and to encourage them to set a challenging goal, we provided goal-setting guidelines based on usage histories. Half of the participants followed the recommendation, but the other half didn't. Such guidance in goal setting can prevent the users from setting overly loose goals, because setting proper or even slightly higher goals would lead to better goal achievement rates [44]. Because of the high diversity of users with different smartphone usage patterns, it still remains a challenge how the goals needs to be appropriated to each user. However, it is evident that regardless of the lockout intensity, setting a goal that is achievable is important, suggesting our proposed approach of 10-20% reduction can be desirable. It is also important to restrict the upper limit of time limit goals. Therefore, for goal setting, we can systematically guide the user into the default setting that gradually reduces the time limit goal, and this goal can be flexible only up to a certain limit strictly controlled by the strong-ILM.

6.1.3 Leveraging Lockout Awareness and Configuring Personalized Reset Times. Smartphone usage times were significantly reduced with restricted lockouts as opposed to permissive warnings. Our exit interview revealed notable behavioral evidence of gradual tension building with the use of stronger ILMs toward the point of the lockout event. This includes frequently checking the remaining time, and having short usage sessions. Therefore, the likelihood of being locked out in the near future served as stimuli to increase awareness of the goal and to engage more in proactive planning behavior of smartphone usage. In contrast, without the lockout, awareness of their goal was low and even ignored. This type of "anticipated punishment" acts as a proximal microboundary [16] that provides continual friction with regard to the user interactions.

Behaviors after the lockout varied across users. A majority expressed constant frustration and pressure from being locked out. However, some expressed a sense of relief due to the lockout. The major difference mainly came from the "remaining time until the reset." In our design of the strong-ILM version, the lockout was reset at midnight. Most participants did not care much about being locked out if it occurred several hours before midnight. If the phone was locked too early, they tended to complain about its restrictiveness. Interestingly, we observed that users occasionally waited until midnight and used their phones afterward, which could not only lead to overuse but also disturb their sleep. Furthermore, users may complain about a lack of overall usage time during the next day. Overall, this observation implies that *reset timing* is also an important design dimension that should be carefully set by considering an individual's circadian rhythm [1].

6.2 Flexible Design for Acceptability of ILMs

6.2.1 Flexibility with Initial Goal Setting. We expected that some participants would opt out due to the strong lockout intensity that they would face in this study. However, no participants opted out of the three-week experimental period. One of the main contributing factors was that we gave them the authority to set their own time limit goals. The results showed that participants allocated longer time limit goals with stronger ILMs and shorter time limit goals for lighter ILMs. This tendency indicates they flexibly controlled the total lockout intensity using the time limit goal. The total lockout intensity of GoalKeeper can be expressed as *ILM Intensity* × *Time Limit Goal*. Despite the varying levels of perceived coercion and frustration across the three versions of GoalKeeper, we can state that the total lockout intensity was maintained within each individual's acceptance boundary due to this flexibility.

Another contributing feature was the one-time chance of user goal modification. The limited modification opportunity was given to make their goal times during their use appropriate, but without compromising the appropriate level of the lockout intensity. This helped the participants to adjust their goal times to meet their out-of-routine requirements. It is important to note that overly flexible goal modifications could lead to decreased effectiveness of the ILMs. Multiple goal modifications when encountering impulsive situations are likely to undermine one's initial goal setting for behavioral changes.

6.2.2 Flexibility for Out-of-routine Contexts. Only five participants modified their goals with the non-ILM version, but 20 did so with the strong-ILM version, implying that the stronger the ILM becomes, the more the need for flexibility increases. The weekday goal extension times did not differ much across the conditions, but the differences between the weekend goals, particularly between the non-ILM and strong-ILM conditions, were large.

We asked when and why the participants modified their goals. The majority replied that there were in *out-of-routine* contexts, such as going on a one-day business trip to another city, visiting a friend or a relative, or participating in an activity outside of their workspace. In some of these contexts, the excessive use of the smartphone was somewhat unavoidable. The modification time in Figure 7 shows a greater time extension with strong-ILM than with non-ILM on weekends. This implies that 1) the concern over the inability to cope with out-of-routine contexts forced them to set a safe yet loose goal, and 2) these out-of-routine contexts mainly occurred on weekends. Prior works have also noted this problem and provided a *temporary unlimit* feature, which was reported to be very useful [30, 37, 38]. This is a user-initiated, temporary allowance. However, if a situational task requires prolonged usage, this approach may be less appropriate. Another approach for managing out-of-routine contexts involves the use of context-aware technology [42]. An example would be location-aware goal setting that can adaptively increase the time limit goal when users are out of their routine workspaces. Another approach would be the use of whitelisting or blacklisting of applications by carefully considering interaction meaningfulness to minimize the impact of a lockout [32, 45]. Context-awareness can be incorporated into whitelisting/blacklisting as well. However, this approach is limited in that there are many neutral applications that could be both productive and counterproductive depending on various contextual factors [31, 34, 45]. Thus, it is important systematically to study out-of-routine context-management strategies that can improve the acceptance of ILM-based applications.

Another reason behind goal modifications was a lack of self-control. Unlike uncontrollable out-of-routine incidents, the participants could have exerted greater self-control to manage their smartphone usage times. However, some participants failed to do so, and extended their time limit goals for entertainment purposes. For such incidents, providing flexibility would be inappropriate, weakening the effectiveness of the lockout mechanism. Therefore, more accurate detection of a user's use intentions [45] or leveraging social support [27, 36] and commitments [51] would be crucial to mitigating such problems.

6.2.3 Adaptability to Lockout Objectives. Smartphone users can have different objectives when using an interaction lockout strategy, and these objectives can change over time. For example, our study results show that there are significant differences in usage limits between weekdays and weekends. Moreover, in a survey on preferences, many participants noted the possible helpfulness of strong-ILM when they need to focus on something (e.g., exams), although they preferred this method less in general. Therefore, a smartphone intervention mechanism should be carefully designed after considering the motives of why the user wants to self-limit their usage. The motives can be related to short-term objectives (e.g., upcoming quizzes) or long-term goals (e.g., getting good SAT scores). It will be useful to provide adaptive lockout mechanisms which can consider their short- or long-term objectives and also be adapted to users' current situations.

7 LIMITATIONS AND FUTURE WORK

There are several limitations associated with our work. First, the sample was limited to graduate students, a group which does not represent the general population. However, this group may share common ground with researchers and knowledgeable workers, as their daily activities occur around the workplace. Second, we found evidence that the effects of ILMs may have some dependence on personal traits, such as self-efficacy, self-control, smartphone engagement, and addiction. Previous work suggests that emotional states can also influence smartphone use behaviors [48]. Moreover, boredom proneness by users is a known predictor of smartphone use frequency and app usage patterns [47] and boredom is known to influence attentiveness to smartphones [58]. However, in this current body of work, we did not take these strata into consideration.

Another limitation is that the experimental period lasted only for one week, per condition. This was enough for a comparative observation of the short-term effects during the use of our intervention application. However, it may not be enough to observe the long-term effects, or directly experience ILMs. Particularly, the adaptation process, along with goal setting and adhering to ILMs, must be observed further. This is due to the fact that sensitivity towards the intervention may diminish over time.

Another limitation is the inclusion of the *break button* feature that may have diluted the intervention effect. We intended to observe how frequently self-triggered temporal breaks were used in the "goal-oriented work" contexts. Our result showed that it was less frequently used, and their impact was limited when compared with that of ILMs. We would have observed greater effects of the ILM interventions without the break mode. However, given that the break mode was given across all experimental conditions, and its usage frequency was fairly low. For these reasons, we think that the break mode did not seriously compromise our results.

Lastly, we found several cases of technology workarounds with four participants that occurred due to software bugs. The participants reported that the bugs were found by chance. One enabled the "energy saving mode," and it stopped GoalKeeper's timer. Another participant could use the phone during the lockout using the "dual-window mode." Two other participants found that when using the phone call during the lockout, GoalKeeper would not come back to the lockout screen. All of these cases were technical issues dependent on certain Android versions or smartphone brands. However, it is interesting to note that once these bugs were found, the self-regulatory commitment level of the users appeared to be weakened and they continued to use their phones. A few participants attempted to find workarounds during the lockout, but failed to do so, soon accepting their situations.

Three participants were eliminated to preserve the validity of the user data. Previous bodies of literature adopting coercive and restrictive interventions have reported similar technology workarounds at vehicles [17], and in work environments [31]. Considering that this issue has been frequently observed in other studies as well, it is worthwhile to explore such workarounds in a deeper level to understand how they influence the intended intervention and, how they can be better managed.

8 CONCLUSION

We studied the interaction lockout mechanisms (ILMs) to understand their effects on reducing smartphone usage in comparative ways. The quantitative results show that these mechanisms significantly reduced the use of smartphones compared to the baseline, a stronger lockout led to a greater reduction in smartphone use. However, the strong-ILM version was least preferred due to the pressure and frustration caused by possible inconvenience during the lockout period. We found that a light degree of ILM was the most balanced approach with reasonable user acceptance as well as effectiveness. We discuss the variations of user experiences with each of the ILMs, and provide insights into the design of ILM-based interventions moving forward.

At this current stage, we have only uncovered how different types of ILMs are utilized for regulating smartphone use. The proposed ILMs can be extended to inhibit any undesired interaction with interactive systems. For example, if a user decides to restrict their TV viewing time, the TV may stop functioning when the user exceeds their time goal. Such single interactions can also be extended to multi-device environments where multiple devices can be leveraged to with regard to a user's undesired interaction and promote positive behaviors. For example, if one does not fulfill a self-defined daily exercise goal, a smartwatch may notify a refrigerator not to open. As ubiquitous computing environments mature, more proactive ILMs can be designed with a broader spectrum of computing devices to facilitate positive, healthier, and safer interactions.

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 (NRF-2017M3C4A7083534), by Institute for Information communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. R7124-16-0004, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding), and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2018R1C1B5044211).

REFERENCES

[1] Saeed Abdullah, Mark Matthews, Elizabeth L. Murnane, Geri Gay, and Tanzeem Choudhury. 2014. Towards Circadian Computing: "Early to Bed and Early to Rise" Makes Some of Us Unhealthy and Sleep Deprived. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 673–684.

[2] Saeed Abdullah, Elizabeth L Murnane, Mark Matthews, Matthew Kay, Julie A Kientz, Geri Gay, and Tanzeem Choudhury. 2016. Cognitive rhythms: Unobtrusive and continuous sensing of alertness using a mobile phone. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 178–189.

[3] National Highway Traffic Safety Administration et al. 2012. Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. *Washington, DC: National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT)* (2012).

[4] Ritu Agarwal and Elena Karahanna. 2000. Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS quarterly* (2000), 665–694.

[5] Icek Ajzen. 1991. The theory of planned behavior. *Organizational behavior and human decision processes* 50, 2 (1991), 179–211.

[6] Freedom App. 2018. Website. https://freedom.to Accessed: 2018-11-12.

[7] David H Barlow and Steven C Hayes. 1979. Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis* 12, 2 (1979), 199–210.
[8] Roy F Baumeister, Kathleen D Vohs, and Dianne M Tice. 2007. The strength model of self-control. *Current directions in psychological science* 16, 6 (2007), 351–355.
[9] Beeminder. 2018. Website. https://www.beeminder.com Accessed: 2018-11-12.
[10] 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, 2005–2014.
[11] Korea Bizwire. 2018. New App Combats Smartphone Addiction Among Youth. http://koreabizwire.com/new-app-combats-smartphone-addiction-among-youth/118541 Accessed: 2018-11-12.
[12] Barry W Boehm, MJ t Seven, and RA Watson. 1971. Interactive problem-solving: an experimental study of lockout effects. In *Proceedings of the May 18-20, 1971, spring joint computer conference*. ACM, 205–210.
[13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[14] Gharad Bryan, Dean Karlan, and Scott Nelson. 2010. Commitment Devices. *Annual Review of Economics* 2, 1 (2010), 671–698.
[15] Sunny Consolvo, David W McDonald, Tammy Toscos, Mike Y Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, et al. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1797–1806.
[16] 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, 1389–1397.
[17] Janet I Creaser, Christopher J Edwards, Nichole L Morris, and Max Donath. 2015. Are cellular phone blocking applications effective for novice teen drivers? *Journal of safety research* 54 (2015), 75–e29.
[18] Laura Dabbish, Gloria Mark, and Víctor M González. 2011. Why do i keep interrupting myself?: environment, habit and self-interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3127–3130.
[19] Tao Dong, Mark W Newman, Mark S Ackerman, and Sarita Schoenebeck. 2015. Supporting reflection through play: field testing the home trivia system. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 251–262.
[20] Brian J Fogg. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002, December (2002), 5.
[21] Drew Fudenberg and David K Levine. 2006. A dual-self model of impulse control. *American economic review* 96, 5 (2006), 1449–1476.
[22] Arup Kumar Ghosh, Karla Badillo-Urquiola, Shion Guha, Joseph J. LaViola Jr, and Pamela J. Wisniewski. 2018. Safety vs. Surveillance: What Children Have to Say About Mobile Apps for Parental Control. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 124, 14 pages.
[23] Scott D Halpern, David A Asch, and Kevin G Volpp. 2012. Commitment contracts as a way to health. *Bmj* 344, 1 (2012), e522–e522.
[24] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
[25] Alexis Hiniker, Sharon S Heung, Sungsoo Ray Hong, and Julie A Kientz. 2018. Coco's Videos: An Empirical Investigation of Video-Player Design Features and Children's Media Use. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 254.
[26] Alexis Hiniker, Sungsoo Ray 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, 4746–4757.
[27] 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 (IDC '17)*. ACM, New York, NY, USA, 85–95.
[28] Esra Erkol İNal, Kadir Demirci, Azİze ÇetİntÜrk, Mehmet Akgönül, and Serpİl Savaş. 2015. Effects of smartphone overuse on hand function, pinch strength, and the median nerve. *Muscle & nerve* 52, 2 (2015), 183–188.
[29] Auk Kim, Woohyeok Choi, Jungmi Park, Kyeyoon Kim, and Uichin Lee. 2018. Interrupting Drivers for Interactions: Predicting Opportune Moments for In-vehicle Proactive Auditory-verbal Tasks. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 175 (Dec. 2018), 28 pages. https://doi.org/10.1145/3287053
[30] 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), 63.
[31] 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), 64.
[32] 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 (CHI '19)*.
[33] Young-Ho Kim, Eun Kyoung Choe, Bongshin Lee, and Jinwook Seo. 2019. Understanding Personal Productivity: How Knowledge Workers Define, Evaluate, and Reflect on Their Productivity.. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*.
[34] Young-Ho Kim, Jae Ho Jeon, Eun Kyoung Choe, Bongshin Lee, KwonHyun Kim, and Jinwook Seo. 2016. TimeAware: Leveraging framing effects to enhance personal productivity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM,

272–283.

[35] Sheila G Klauer, Feng Guo, Bruce G Simons-Morton, Marie Claude Ouimet, Suzanne E Lee, and Thomas A Dingus. 2014. Distracted driving and risk of road crashes among novice and experienced drivers. *New England journal of medicine* 370, 1 (2014), 54–59.

[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, 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, 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, 1235–1245.

[39] Min Kwon, Dai-Jin Kim, Hyun Cho, and Soo Yang. 2013. The smartphone addiction scale: development and validation of a short version for adolescents. *PloS one* 8, 12 (2013), e83558.

[40] Heyoung Lee, Heejune Ahn, Samwook Choi, and Wanbok Choi. 2014. The SAMS: Smartphone addiction management system and verification. *Journal of medical systems* 38, 1 (2014), 1.

[41] Moon-Hwan Lee, Yea-Kyung Row, Oosung Son, Uichin Lee, Jaejeung Kim, Jungi Jeong, Seungryoul Maeng, and Tek-Jin Nam. 2018. Flower-Pop: Facilitating Casual Group Conversations With Multiple Mobile Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 150 (Jan. 2018), 24 pages.

[42] 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.

[43] 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, 43.

[44] Edwin A Locke. 1968. Toward a theory of task motivation and incentives. *Organizational behavior and human performance* 3, 2 (1968), 157–189.

[45] Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. 2018. What Makes Smartphone Use Meaningful or Meaningless? *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 22 (March 2018), 26 pages.

[46] ManicTime. 2018. Website. http://www.manictime.com Accessed: 2018-11-12.

[47] Aleksandar Matic, Martin Pielot, and Nuria Oliver. 2015. Boredom-computer interaction: Boredom proneness and the use of smartphone. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 837–841.

[48] Abhinav Mehrotra, Fani Tsapeli, Robert Hendley, and Mirco Musolesi. 2017. MyTraces: Investigating Correlation and Causation between Users' Emotional States and Mobile Phone Interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 83.

[49] Katherine L Milkman, Julia A Minson, and Kevin GM Volpp. 2013. Holding the Hunger Games hostage at the gym: An evaluation of temptation bundling. *Management science* 60, 2 (2013), 283–299.

[50] Katherine L Milkman, Todd Rogers, and Max H Bazerman. 2008. Harnessing Our Inner Angels and Demons: What We Have Learned About Want/Should Conflicts and How That Knowledge Can Help Us Reduce Short-Sighted Decision Making. *Perspectives on Psychological Science* 3, 4 (2008), 324–338.

[51] Sean A. Munson, Erin Krupka, Caroline Richardson, and Paul Resnick. 2015. Effects of Public Commitments and Accountability in a Technology-Supported Physical Activity Intervention. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1135–1144.

[52] Karla Klein Murdock. 2013. Texting while stressed: Implications for students' burnout, sleep, and well-being. *Psychology of Popular Media Culture* 2, 4 (2013), 207.

[53] Harri Oinas-Kukkonen and Marja Harjumaa. 2009. Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems* 24, 1 (2009), 28.

[54] 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 (MobileHCI '18)*. ACM, New York, NY, USA, Article 4, 12 pages.

[55] Gadget Guide Online. 2016. Changes on Waiting Time for Wrong Input on Galaxy S6 Lock Screen in Android Marshmallow Update. https://tinyurl.com/y84edzkk Accessed: 2018-11-12.

[56] Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing* 16, 1 (2012), 105–114.

[57] Joonyoung Park, Jin Yong Sim, Jaejeung Kim, Mun Yong Yi, and Uichin Lee. 2018. Interaction Restraint: Enforcing Adaptive Cognitive Tasks to Restrain Problematic User Interaction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, LBW559.

[58] Martin Pielot, Tilman Dingler, Jose San Pedro, and Nuria Oliver. 2015. When attention is not scarce-detecting boredom from mobile phone usage. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 825–836.

[59] Martin Pielot and Luz Rello. 2017. Productive, anxious, lonely: 24 hours without push notifications. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 11.

[60] James O Prochaska. 2013. Transtheoretical model of behavior change. In *Encyclopedia of behavioral medicine*. Springer, 1997–2000.

[61] Jun Rekimoto and Hitomi Tsujita. 2014. Inconvenient interactions: an alternative interaction design approach to enrich our daily activities. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. ACM, 225–228.

[62] RescueTime. 2018. Website. http://www.rescuetime.com Accessed: 2018-11-12.

[63] Todd Rogers, Katherine L Milkman, and Kevin G Volpp. 2014. Commitment devices: using initiatives to change behavior. *JaMa* 311, 20 (2014), 2065–2066.

[64] Heather Royer, Mark Stehr, and Justin Sydnor. 2015. Incentives, Commitments, and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company. *American Economic Journal: Applied Economics* 7, 3 (July 2015), 51–84.

[65] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[66] Kathleen Ann Sheehan and Tom Burns. 2011. Perceived coercion and the therapeutic relationship: a neglected association? *Psychiatric Services* 62, 5 (2011), 471–476.

[67] In-geon Shin, Jin-min Seok, and Youn-kyung Lim. 2018. Too Close and Crowded: Understanding Stress on Mobile Instant Messengers Based on Proxemics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 615, 12 pages.

[68] Jaemyung Shin, Bumsoo Kang, Taiwoo Park, Jina Huh, Jinhan Kim, and Junehwa Song. 2016. BeUpright: Posture Correction Using Relational Norm Intervention. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 6040–6052.

[69] stickK. 2018. Website. http://www.stickk.com Accessed: 2018-11-12.

[70] Nynke Tromp, Paul Hekkert, and Peter-Paul Verbeek. 2011. Design for socially responsible behavior: a classification of influence based on intended user experience. *Design Issues* 27, 3 (2011), 3–19.

Received May 2018; revised November 2018; accepted January 2019