

# StayFocused: Examining the Effects of Reflective Prompts and Chatbot Support on Compulsive Smartphone Use

Zhuoyang Li\*

Department of Computer Science  
City University of Hong Kong  
Hong Kong, China  
zhuoyanli4@cityu.edu.hk

Ray LC

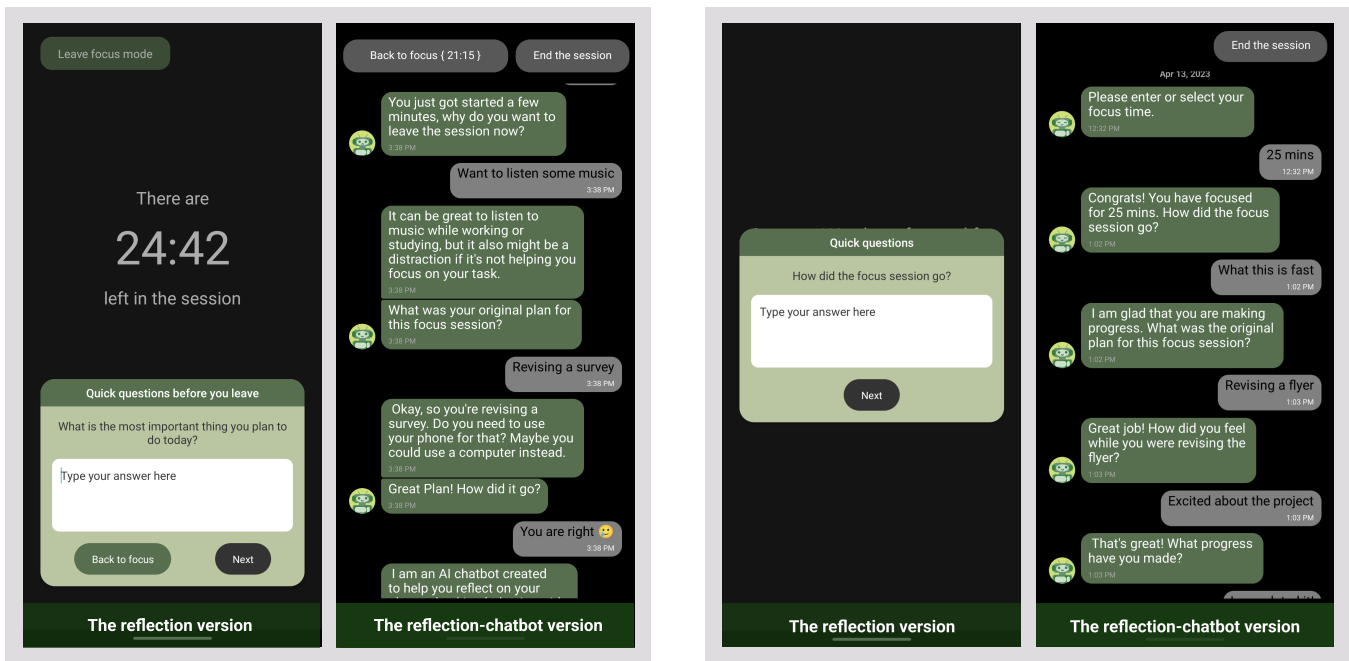
School of Creative Media  
City University of Hong Kong  
Hong Kong, China  
recfreq@gmail.com

Minhui Liang\*

Department of Computer Science  
City University of Hong Kong  
Hong Kong, China  
mhliang4-c@my.cityu.edu.hk

Yuhan Luo

Department of Computer Science  
City University of Hong Kong  
Hong Kong, China  
yuhanluo@cityu.edu.hk



(a) Scenario 1: Leaving a focus session earlier.

(b) Scenario 2: Completing a focus session.

**Figure 1: The reflection and reflection-chatbot versions of StayFocused, which sends reflective prompts to people whenever they (a) attempt to leave a focus session earlier before it ends or (b) complete a focus session. In the reflection version, people will be prompted to answer predefined questions about their phone-checking intentions; in the reflection-chatbot version, they will answer a similar set of questions led by a chatbot.**

\*Equal contribution.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
CHI '24, May 11–16, 2024, Honolulu, HI, USA  
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 979-8-4007-0330-0/24/05...\$15.00  
<https://doi.org/10.1145/3613904.3642479>

## ABSTRACT

Amidst the increasingly prevalent smartphone addiction, we introduce StayFocused, a mobile app to help people focus on their tasks at hand by reducing compulsive smartphone use. Besides guiding people to set focus sessions for non-screen time, we incorporated reflective prompts probing individuals' phone-checking intentions whenever they check their phones and a chatbot to deliver these prompts. To examine the effects of the reflective prompts and the chatbot support, we designed three versions of StayFocused: baseline, reflection, and reflection-chatbot, and conducted a stage-based

between-subjects study with 36 college students over five weeks. We found that participants who received the reflective prompts were able to focus longer and resist distractions, and those with chatbot support seemed to better maintain their smartphone use reduction. By highlighting how participants reflected on their focus session activities and their preferences for the chatbot, we discuss the implications of designing persuasive conversational interfaces to reduce unintended behaviors.

## CCS CONCEPTS

• **Human-centered computing** → **HCI design and evaluation methods**; **User interface design**; *Natural language interfaces*.

## KEYWORDS

smartphone addiction, reflection, persuasive technology (PT), conversational UI, large language model

### ACM Reference Format:

Zhuoyang Li, Minhui Liang, Ray LC, and Yuhan Luo. 2024. StayFocused: Examining the Effects of Reflective Prompts and Chatbot Support on Compulsive Smartphone Use. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3613904.3642479>

## 1 INTRODUCTION

Smartphones have become an integral part of our daily lives, offering instant connectivity and access to vast information. On the other hand, there is a growing concern regarding compulsive smartphone use, a behavior pattern resulted from excessive reliance on smartphones, which involves repetitive and uncontrollable phone checking despite being aware of the associated negative impacts [15, 48]. Compulsive smartphone use has been observed across different age groups, degrading people’s physical health, mental well-being, and work productivity [16, 17, 20, 48, 56, 74]. Younger generations, who are exposed to smartphones at an early age, are particularly susceptible to compulsive smartphone use [30]. In the United States, 15% of young adults between 18 and 29 years old are found heavily dependent on smartphones [70]. For example, college students spend more than five hours a day on their phones, even during classes where smartphones are prohibited [34]. This habit can cause distractions and interruptions to students’ work routines, leading to declined ability to focus, and ultimately reduced academic performance and cognitive functioning [66].

To raise people’s awareness of their smartphone use behaviors, mobile operating systems have made digital well-being a built-in feature through monitoring screen time and setting time limit for distracting apps. However, a recent study highlighted that overcoming the inherent fear of missing out (FOMO) poses a challenge for people who feel anxious about being disconnected from social media and news updates [21]. Thus, solely relying on screen time monitoring or restricting the use of certain apps may not help them truly reflect on how and why they spend time on smartphones [4, 17]. Taking a different approach, researchers have designed and developed productivity tools that guide individuals to set dedicated focus (non-screen) time and incorporated rewarding mechanisms such as symbolic visualizations of one’s focus progress (e.g., Forest [22], MyTime [31]). While these productivity tools could help reduce

overall screen time [31], it is unclear whether the induced behavior change can sustain if people stop using the tools. Additionally, there has been a limited understanding of people’s phone use behavior during the designated focus time, such as why check their phones and how this behavior affects their subsequent plans and actions.

In this work, we aim to help individuals stay focused on their tasks at hand when smartphone use is unnecessary. That is, minimizing their smartphone use as much as possible during the intended focus time. We designed and developed StayFocused, a mobile app that enables people to plan for focus sessions by setting a duration for *not* interacting with their phones. Whenever people check their phones during the focus session, StayFocused will send *reflective prompts* asking about the reasons behind their phone-checking intentions and questions related to their original plans (see Figure 1a). This design idea is partly drawn from the concept of *situated reflection*—probing individuals to articulate their intentions in a specific situation where the target behavior occurs [67], which was found promising in inducing immediate behavior change [10, 49]. Moreover, inspired by prior work that leveraged a chatbot to facilitate self-reflection on physical health and work activities [41, 42], we incorporated a chatbot to deliver the reflective prompts and interact with people in natural languages. We envision the chatbot can further enhance the effects of the reflective prompts and increase individuals’ commitment to their focus plans.

To examine the effects of the *reflective prompts* and the combination of the *reflective prompts* plus *chatbot support*, we designed and developed three versions of StayFocused:

- The *baseline* version, which allows individuals to start a focus session by choosing a duration for *not* interacting with their phones. Individuals could leave the session earlier before it ends simply by clicking the button “Leave focus mode” and confirming the decision.
- The *reflection* version, which is built upon the baseline version with reflective prompts asking individuals to explain their situation (in a textbox) whenever they attempt to leave the focus session and complete a session.
- The *reflection-chatbot* version, which is built upon the reflection version but leverages a chatbot (powered by GPT-3) to deliver the reflective prompts and to provide feedback to individuals’ responses.

We conducted a stage-based between-subjects study by deploying StayFocused to 36 college students who self-identified with problematic smartphone use. Participants were assigned into three groups to use different versions of StayFocused for three weeks: the PB group used the baseline version ( $n = 11$ ), the PR group used the reflection version ( $n = 13$ ), and the PRC group used the reflection-chatbot version ( $n = 12$ ). To assess participants’ smartphone use, StayFocused gathered their screen duration before (one week), during (three weeks), and after (one week) using StayFocused. We also employed pre- and post-intervention surveys to collect participants’ problematic smartphone use and self-regulation ability, followed by semi-structured interviews to understand their overall experience with StayFocused.

We found that although the three groups did not differ in their focus session completion rate, the PR and PRC groups tended to focus longer than the PB group. During the intervention stage, all three groups exhibited a reduction in their screen duration; in the

post-intervention stage, only the PRC group further reduced their screen duration, although the reduction is not statistically significant. The interview findings and participants' responses to the reflective prompts further complemented the quantitative results, showing that the reflective prompts enabled participants to reflect on their smartphone dependence and encouraged them to resist distractions from their phones. The chatbot was found to increase PRC participants' accountability and offered emotional support by acknowledging their efforts to stay focused. Despite the promises, we noted the limitations of the chatbot and the counterproductive consequences it may introduce. With the lessons learned, we discussed what affected participants' phone-checking behavior during intended focus time, and how to better integrate reflective prompts and chatbot support to reduce compulsive smartphone use.

To the best of our knowledge, this work is among the first to explore the opportunity for incorporating reflective prompts and chatbot support to mitigate unintended behaviors. Our findings contributed to the Human-Computer Interaction (HCI) community in three folds: (1) an empirical understanding of how reflective prompts, delivered from a traditional dialogue interface *versus* a chatbot, enabled individuals' reflection *in situ* and affected their subsequent behaviors; (2) an in-depth understanding of the nuances associated with individuals' preferences for the reflective prompts and the chatbot, and how these nuances played parts in their overall experiences in the context of reducing compulsive smartphone use; and (3) design implications for better incorporating reflective prompts and chatbots to promote sustainable behavior change.

## 2 RELATED WORK

### 2.1 Compulsive Smartphone Use and Existing Interventions

Compulsive smartphone use, a behavior pattern arising from excessive reliance on smartphones, often involves repetitive and uncontrollable phone-checking, even when the individual is aware of the negative impacts associated with such behaviors [14, 15, 48, 56]. Research has shown that compulsive smartphone use often occurs during times of stress or as a form of procrastination [2, 24], which can significantly disrupt one's work routines and other daily activities, and contributes to various physical and mental health problems, including poor sleep quality, reduced cognitive functioning, depression, and anxiety [15, 56]. Younger generations, particularly college students, are susceptible to compulsive smartphone use [3, 54]. Partly, modern smartphone applications are designed with highly addictive features such as infinite scrolling, push notifications, and personalized advertisements that are difficult to disengage [17]. Individual life circumstances, such as emotional difficulties and loneliness can further exacerbate the situation [17, 44, 46, 80]. Also, younger individuals tend to have weaker self-regulation abilities for time management and resist such technological temptations [16], especially in the early stages of emancipation from parental oversight (e.g., fresh college students).

Interventions to help people better manage their smartphone use behaviors have been developed by phone manufacturers, academic researchers, and third-party companies. These interventions

employed strategies including monitoring phone usage (e.g., Digital Wellbeing [6]), blocking distracted applications (e.g., Let's FOCUS [36]), setting goals for non-screen time (e.g., Forest [22], MyTime [31]), increasing the difficulties to unlock screens (e.g., LocknType [37]), and parental regulations [16]. Although there is evidence showing the effectiveness of these interventions in reducing smartphone use (e.g., shorter screen time [31], less unlocks [37]), those with limited self-regulation abilities may disregard the restriction once they figure out how to bypass it [15]. In the case of Android Digital Wellbeing, for example, people can still open the blocked apps simply by ignoring or removing the restriction. More importantly, existing work rarely investigated whether the induced behavior changes in smartphone use can sustain after withdrawing the intervention. As suggested by psychological research, restricting compulsive behavior can lead to a counterproductive effect, causing people to relapse to their previous habits with even stronger compulsion once the restriction has been withdrawn [62]. In this light, our work not only examines the effects of reflective prompts and chatbot support on smartphone use, but also explores how sustainable the effects are.

### 2.2 Enabling Reflection for Positive Behavior Change

To encourage positive behavior change, a growing body of research has emphasized the importance of self-reflection—an activity of critically thinking about one's feelings and behaviors, along with the underlying reasons. Across domains (e.g., healthy eating, exercise, work productivity), researchers have attempted to enable "reflection-on-action" by aggregating one's behavior data across time and designing visualizations to help them understand their behavior patterns [53, 58, 59, 65]. For instance, Rooksby et al. designed and built ScreenLife, which helped people develop awareness of their multi-device usage and increased their motivation to reduce unnecessary digital interaction [65]; Kim et al. developed TimeAware and compared how different visual framings of productivity scores supported knowledge workers to reflect on their work performance and improve their productivity [38]. Reflection-on-action, therefore, allows people to explore and learn about their past behaviors, so that they can use this self-knowledge to improve their behaviors in the future.

In the meantime, previous work has shown that situated reflection on one's current behavior has the potential to introduce immediate behavior change [42, 49, 52]. This type of reflection is often activated by prompting people to articulate their intentions and thoughts when the target behavior occurs, which allows them to critically think about and assess their behaviors *in situ* [33, 49]. For example, Luo et al. designed FoodScrap, a food journaling app asking people to explain their food decisions close to the time of eating, such as why they decided to choose certain foods and when to eat [49]. These questions prompted people to reflect on the healthiness of their choice and even reconsider their food decisions [49]. In a similar vein, Terzimehić et al. developed MindPhone to promote mindful smartphone use with two prompts delivering when people unlock their phones: one focusing on the reasons behind one's phone-checking behavior and the other focusing on their subsequent plans [71]. Through a real-world deployment study, the researchers found that the latter prompt significantly reduced absentminded phone use [71]. Inspired by this study, we created a

set of “*reflective prompts*” that help individuals reflect on why they check their phones during the non-screen time they had planned, and envision that the situated reflection can prevent them from compulsively checking their phones when it is unnecessary. Besides supporting situated reflection when people are distracted, we also aim to combine reflection-on-action when they achieve their focus goals. Therefore, our reflective prompts include questions related to individuals’ phone-checking intentions as well as their feelings upon completing a focus session. More importantly, we go beyond this work by incorporating a chatbot to deliver these prompts and examine its effects on individuals’ behaviors.

### 2.3 Chatbot As a Persuasive Agent

Chatbots, also known as conversational agents, have been widely used to engage people in natural language conversations, ranging from online tutoring and customer service to even psychological counseling [25, 75]. The anthropomorphic features of chatbots (e.g., identity cues, personality) hold great potential to deliver persuasive messages aimed at changing people’s beliefs, attitude, and behaviors [9, 57, 75]. Research showed that people tend to attribute social roles to the chatbots they interact with [57], and are more receptive to the messages sent by the chatbot compared with plain text [9]. In the HCI and health community, chatbots have been integrated into various applications that aim to enable reflection and behavior change regarding health-related goals [5, 26, 40–42, 50, 79]. For example, Kocielnik et al. built a virtual reflection companion that talks to people about their physical activities, helping them summarize their previous data and asking them to explore their activity patterns as a way to encourage more exercise [42]. Through a field study, the researchers found that these mini-dialogues between the reflection companion and individuals successfully enabled in-depth reflection and led to the adoption of new exercise behaviors [42]. Likewise, Alphonse et al. investigated how a chatbot served as a virtual coach to support smoking cessation; they found that the perceived human likeness of the chatbot increased people’s accountability towards their goals and effectively helped them maintain their progress with relevant feedback and advice [5].

Mostly relatedly, Grover et al. designed and built a chatbot to help people schedule and block time slots on their calendars, so that they could plan and focus on important tasks [26]. By adding a video avatar and emotion recognition ability to the chatbot, the researchers found that participants were likely to schedule more tasks and become more productive [26]. While this work investigated people’s task-planning behaviors with the support of a chatbot, we are interested in examining the persuasiveness of a chatbot in helping people reflect on and reduce phone-checking behaviors during intended focus time. Furthermore, the advancement in natural language processing (NLP), especially the recent surge of LLMs, has opened up exciting opportunities for designers and developers to customize chatbots that engage people in more natural and fluent conversations [1, 28, 29, 76]. However, little work has applied LLMs to build a chatbot in the context of enabling situated reflection and reducing unintended behaviors. Our work, therefore, takes the first step to exploring the design opportunities.

## 3 STAYFOCUSED

This work aims to answer two research questions in a situation where an individual plans to focus on a task without smartphone distraction: (1) whether and how can incorporating reflective prompts help them reflect on their phone-checking behaviors and reduce their overall smartphone use? (2) whether and how can leveraging a chatbot to deliver the reflective prompts further strengthen the abovementioned effects? Here, we describe our design goals to answer the two research questions, as well as the design components of StayFocused and implementation details.

### 3.1 Design Goals

**3.1.1 DG1. Enabling Situated Reflection and Reflection-on-Action.** As we mentioned in Related Work, existing interventions on reducing smartphone use largely focused on supporting individuals to reflect on their past phone usage (i.e., *reflection-on-action* [6, 22, 31, 65, 67]). Only one study (i.e., MindPhone [71]) enabled individuals to reflect on their phone-checking behavior in real-time (i.e., *situated reflection*). Our work, on the other hand, aims to support both situated reflection during the moment when individuals are checking their phones and reflection-on-action after they have completed a focus session without phone distraction. We envision that situated reflection can induce immediate change in one’s phone-checking intention, while reflection-on-action can help sustain the behavior change by offering individuals a sense of achievement. Therefore, we need to design two different sets of reflective prompts tailored to these two scenarios.

**3.1.2 DG2. Leveraging a Chatbot As a Persuasive Agent.** Prior research has shown that a chatbot equipped with anthropomorphic attributes such as identity cues [25, 75] and advanced natural language understanding abilities [79], has the potential to serve as a persuasive agent for health lifestyle promotion. In addition, tailoring the chatbot’s responses to people’s emotions can create a sense of support and encourage continued use [33, 35, 69]. Thus, it is important to design a natural, coherent, and personalized conversation experience, which necessitates the utilization of state-of-the-art language models.

**3.1.3 DG3. Mitigating Fatigue Effects While Preventing Over-Engagement.** As an application that supports daily reflection on one’s smartphone use, it is important to mitigate fatigue effects (e.g., avoid people from being desensitized to reflective prompts [4, 49]). Thus, we recognize the importance of reducing the repetitiveness of the prompts to ensure that individuals remain alert and engaged throughout the conversation [61]. At the same time, we need to prevent individuals from *over-engaging* with the app, especially the version with a chatbot, because our ultimate goal is to reduce smartphone use when one’s priority is to focus on other tasks. In this regard, we need to center the conversation topics around individuals’ focus plans and their phone-checking intentions, and set a limit for conversation turnarounds.

**3.1.4 DG4. Supporting Research Exploration.** To answer our research questions on the effects of reflective prompts and chatbot support in helping people reflect on their compulsive smartphone use and reduce this behavior, we need to isolate and analyze the reflective prompts and chatbot support as independent variables. This requires us to develop different versions of StayFocused—one does not include any reflective prompts and chatbot support, one

includes only the reflective prompts, and one includes both the reflective prompts and chatbot support, respectively. This strategy has been applied in previous work that compared the effects of different design elements on people’s behaviors [38, 50]. Although the ways individuals interact with regular prompts and chatbot-delivered prompts may differ by nature, we aim to deliver the same set of prompts to minimize confounding factors.

## 3.2 StayFocused Interaction Flow

The following describes the interaction flow with the three versions of StayFocused (see Figure 2), and how this process addresses the above design goals.

**3.2.1 Starting a Focus Session.** The three versions of StayFocused (baseline, reflection, reflection-chatbot) share the same mechanism of starting a focus session (DG4). When StayFocused is launched, people are asked to set a focus duration (from 25 to 125 minutes suggested by the Pomodoro technique [18]), during which they plan to avoid using their smartphones (see Figure 3). In the baseline and reflection versions, the duration can be manually entered or selected from a drop-down list. In the reflection-chatbot version, the process is similar but occurs in a text-based conversation with a chatbot. During the focus session, a countdown timer is displayed at the center of the screen in all versions, with a button “Leave focus mode” at the top left. To remind people to use StayFocused every day, the app sends them a daily reminder through the notification channel (“Time to focus!”). It’s worth noting that we did not block the use of other apps or disable notifications during the focus session, because sometimes checking smartphones may be necessary and urgent (e.g., replying to an important message) [46].

**3.2.2 Leaving a Focus Session.** Individuals can leave the focus session earlier before it ends by tapping the “Leave focus mode” button. In the baseline version, they only need to confirm whether to leave or stay. In the other two versions, four reflective prompts will pop up asking why they want to leave (or check their phones) and other related questions (e.g., challenges in staying focused), as a way to enable situated reflection (DG1). In the reflection version, individuals need to type their responses to the prompts in a textbox; and in the reflection-chatbot version, their responses are collected in a natural language conversation with the chatbot (DG2). To reduce the repetitiveness of prompts (DG3), we created a pool of multiple prompts about smartphone use, distraction sources, alternative activities, focus plan, and productivity. When these prompts are delivered, they are randomly selected from the pool and paraphrased. The first prompt is always about phone-checking intention, and examples of other prompts included “Are there any alternative activities you could engage?,” “Which app are you planning to check?,” and “What is the most important thing you plan to do today?”

People can quickly return to the focus session during their attempts to leave; if they are still determined to leave, they are encouraged to answer at least two questions, although it is not required. If people leave StayFocused without tapping the “Leave focus mode” button (e.g., opening another app, going to the home screen), a notification will appear on the top of the screen, reminding them that the focus session will end in 10 seconds if they do not return to the app. In the latter case with the reflection and reflection-chatbot versions, people will be prompted to explain their prior

phone-checking intentions and plans (“Why did you just check your phone?”) when they open StayFocused next time (this process is illustrated in Figure 2).

**3.2.3 Completing a Focus Session.** Upon completing a focus session, the baseline version simply displays a message (“Congrats! You have focused for 25 minutes.”). In the reflection and reflection-chatbot versions, people are prompted to reflect on their experience during the focus session (e.g., “How did the focus session go?”) and think about next steps (e.g., “What is your next plan?,” “Does the complete session bring you closer to your goal today?”) as a way to enable reflection-on-action (DG1). In the reflection version, the prompts are also randomly selected and paraphrased from our pre-created prompt pool; in the reflection-chatbot version, the chatbot guides the conversation by asking a similar set of questions. Table 1 shows an example conversation with the chatbot in the scenario of attempting to leave a focus session and completing a focus session.

## 3.3 Implementation

StayFocused was built by React Native. To retrieve phone usage statistics, we utilized the UsageStatsManager API [73] provided by Android. We summed up the user’s screen duration and unlock times without tracking their usage of individual apps. The interaction logs are stored in Firebase, which adopted role-based access control (RBAC) on Firebase to ensure that the data can only be accessed by the research team.

To implement the conversational interface of the reflection-chatbot version, we used a UI framework called react-native-gifted-chat [64]. The chatbot was powered by GPT-3 with two prompts<sup>1</sup> designed for the two conversation scenarios: leaving the focus mode and completing a focus session, respectively. Learning from the empirical findings in prior work that leveraged GPT-3 to customize a chatbot [45, 76], we created a persona Alan, a “focus companion” and listed the tasks for the two conversation scenarios in two separate prompts, along with a few conversation examples (shots) aiming at supporting reflection on one’s phone-checking behaviors and enabling reflection on one’s complete focus sessions, respectively (see the appendix for the full-length prompts). The detailed prompt design rationale and curation process were described in our early work [47].

## 4 METHOD

We conducted a field deployment study with a between-subjects design to compare the three versions of StayFocused. We also collected the time that participants spent on their smartphones before (pre-intervention, one week), during (intervention, three weeks), and after (post-intervention, one week) using StayFocused, which allowed us to analyze within-subject differences in their smartphone use. The study was approved by the university’s Ethics Review Committee.

### 4.1 Participants

We advertised the study through the university mailing list, Facebook, and Reddit. Our screening survey received 367 responses, and

<sup>1</sup>Note that the term “prompts” here refers to the instructions used to steer the response of GPT-3, which differs from the “reflective prompts” that are questions sent to encourage reflection on one’s smartphone use.

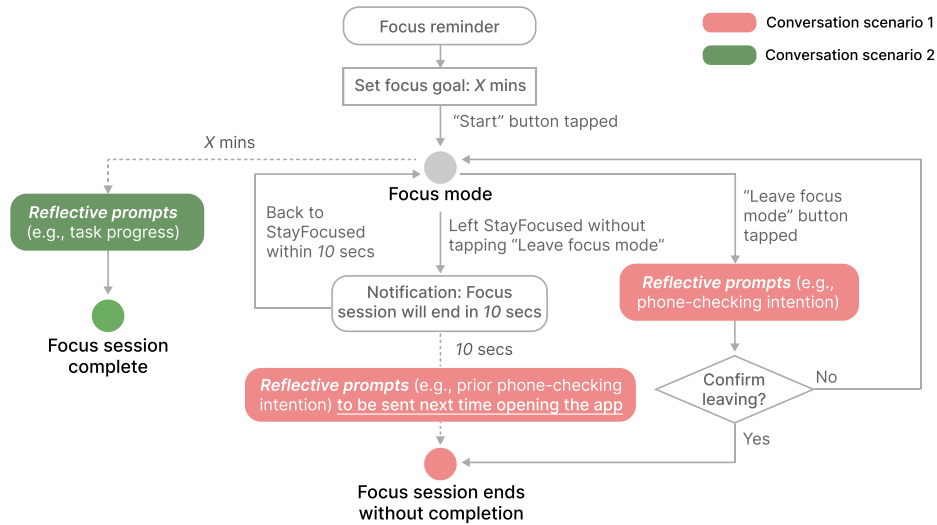
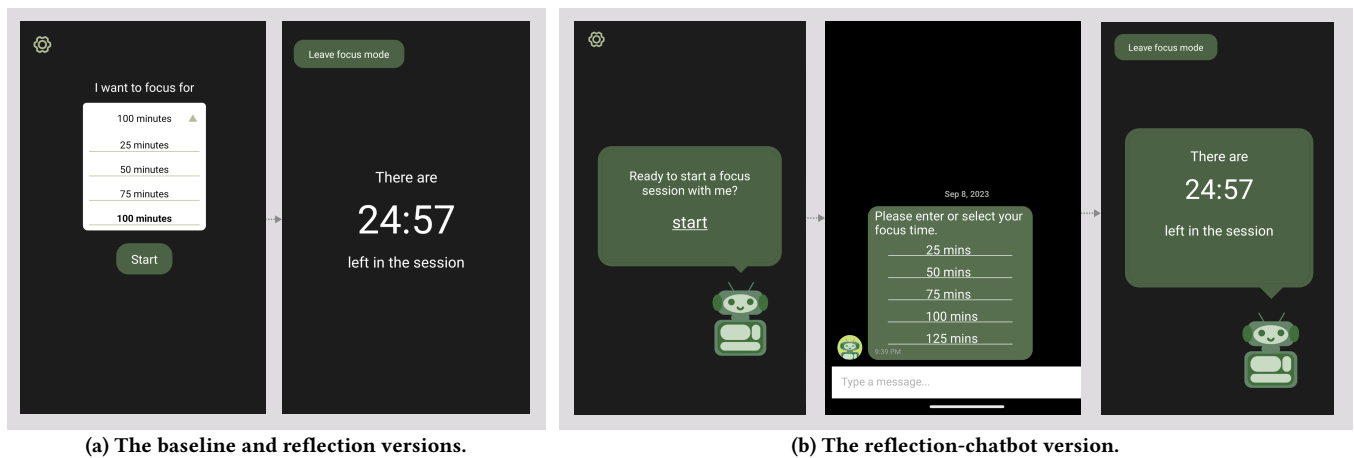


Figure 2: StayFocused interaction flow (The reflective prompts wrapped in the colored boxes are not included in the baseline version).



(a) The baseline and reflection versions.

(b) The reflection-chatbot version.

Figure 3: Setting a focus duration: (a) the baseline and reflection versions share the same interface; and (b) in the reflection-chatbot version, a chatbot facilitates this process through text-based conversations. Upon starting a focus session, a countdown timer is displayed at the center of the screen.

















203 of them met our inclusion criteria: individuals who (1) are over 18 years old; (2) are currently enrolled in a college program; (3) own a smartphone with an Android system (StayFocused only supports Android system); (4) use their smartphones mainly for non-work purposes such as entertainment or social media (i.e., smartphones are not the primary devices for work or study); and (5) are interested in reducing smartphone use and motivated to enhance productivity. Among these individuals, 50 responded to our pre-intervention survey and attended the tutorial session.

Each participant was assigned to one of the three groups: the baseline (PB) group, the reflection (PR) group, and the reflection-chatbot (PRC) group. The group assignment aimed to achieve a balanced distribution of age, gender, and self-reported smartphone use among the participants. As the study progressed, 13 participants dropped out (4 from PB, 5 from PR, and 4 from PRC) because they

stopped using StayFocused and did not reply to our follow-up emails. We also excluded one participant from the PB group due to data loss caused by technical issues.

By the end of the study, 36 participants remained in our analysis: 11 in PB, 13 in PR, and 12 in PRC groups. The participants' ages ranged from 20 to 30 ( $M = 24, SD = 2.34$ ), including 9 females, 26 males, and one non-binary. These participants were also those who expressed a strong motivation to participate in our study during the screening survey. Their motivations include reducing general smartphone use, overcoming social media addiction, improving work and study productivity, learning about their phone use habits, and prioritizing important tasks. The participants were from different regions including the U.S. (30), Hong Kong (4), Kenya (1), and Nigeria (1). All of them were enrolled in a college program and majored

**Table 1: Two conversation examples from our participants who interacted with the chatbot (PRC-5 in scenario 1 and PRC-1 in scenario 2). Our pre-defined prompts are *italicized*.**

Scenario 1: Leaving the focus session earlier	Scenario 2: Completing a focus session
<p> <i>You just got started a few minutes; why do you want to leave the session now?</i></p> <p> I need to switch the song playing in YouTube music.</p>	<p> <i>Congrats! You have focused for 25 minutes. How did the focus session go?</i></p> <p> I rearranged my messy files on the desktop and the laptop.</p>
<p> I can understand wanting to listen to music while you work; but maybe you could try a different platform that isn't as distracting? <i>Are there any alternative activities you could engage in instead of checking your phone?</i></p> <p> What apps do you suggest apart from YouTube music?</p>	<p> Oh; that's great! How many files did you arrange?</p> <p> I can't tell you a concrete number but it makes me feel more organized!</p>
<p> I think any app that can help you relax and focus on your work would be great. Maybe try some white noise or nature sounds apps? <i>What are the challenges during the focus session, if any?</i></p> <p> Do you know what white noise is?</p>	<p> That's great! Being more organized can help you focus better in the future. Do you have any thoughts about focusing for a longer period of time in the future?</p> <p> Yep.</p>
<p> Yes, I do. White noise is a type of sound that is produced by combining different frequencies of sound waves. It is often used to help people focus or relax. <i>How long are you planning to check your phone this time?</i></p> <p> I will get back to the focus session.</p>	<p> Well done! What do you think about the session overall?</p> <p> Maybe I will start another focus session soon after a short break.</p>

in different subjects including Computer Science, Medical Science, Media and Communication, Sociology, Psychology, Biochemistry, and Finance. The majority of the participants were full-time undergraduate students, except for two part-time undergraduate students, one full-time graduate student, and one part-time graduate student. Based on the screening survey, our participants spent 2.69 hours daily on their smartphones for work or study activities and 3.75 hours for non-productive activities. Overall, the three groups of participants did not significantly differ in their gender, age, study major, and self-reported phone use duration.

## 4.2 Study Procedure

The study encompassed six steps: (1) pre-intervention survey; (2) tutorial; (3) three-week intervention deploying StayFocused; (4) post-intervention survey; (5) debriefing interviews; and (6) collection of smartphone use post-intervention. Upon completing the study, we offered each participant a 40 USD Amazon gift card.

**4.2.1 Pre-Intervention Survey.** Before the study, we employed a survey to assess participants' behaviors related to smartphone use. We incorporated the scale of problematic use of mobile phones (PUMP), a commonly used instrument to measure the psychometric properties of problematic phone use [55]. As an example, one of the items in the scale asks participants to indicate their level of agreement with the statement: "*I need more time using my cell phone to feel satisfied than I used to need*," in which participants could rate their agreement on the scale from 1 (strongly agree) to 7 (strongly disagree). The survey also incorporated the Self-Regulation Scale [68], a widely used instrument to assess one's attention control in goal-oriented tasks. This is because prior research showed that addictive smartphone use can be affected by one's self-regulation ability [16].

**4.2.2 Tutorial.** To ensure that participants understand the study procedure, we arranged several tutorials via Zoom. Participants assigned to the same group attended these tutorials together, with fewer than 3 people in each session. During the tutorial, we first introduced the study in brief and then instructed the participants to install the StayFocused app and set up an account. Next, we guided

them to start a focus session and explained different options when they wanted to end the session. We also showed them how to check their focus records.

At the end of the tutorial, we encouraged the participants to (1) regularly use StayFocused for the next three weeks (they were allowed to skip using the app for a maximum of two consecutive days but no more than that); (2) start a focus session whenever they feel a need to refrain from compulsive smartphone use (alternative activities including work, study, relaxation, exercise, etc); and (3) aim for starting four focus sessions per day (completion of the sessions was not mandatory); For the PR and PRC participants, we also encouraged them to honestly answer the reflective prompts. The tutorial session lasted for about 15 minutes for the PB group and about 30 minutes for the PR and PRC groups.

**4.2.3 Intervention Stage: StayFocused Deployment.** Participants started using StayFocused from the day following the tutorial and continued for the next three weeks. Throughout this period, we consistently monitored participants' app usage. To ensure that participants adhered to the study guidelines and in case they encountered any technical issues, we would send them an email reminder if they stopped using StayFocused for over two consecutive days. Upon creating an account on StayFocused, the app automatically gathered participants' smartphone screen duration and unlock frequency over the past 7–10 days, depending on the settings of their phones. During the study, StayFocused also kept collecting these data on a daily basis. At the end of the three weeks, we asked the participants to log out from StayFocused but not to uninstall the app. Thus, StayFocused could keep monitoring their activities on the backend to ensure that they did not start any focus sessions thereafter for an additional week.

**4.2.4 Post-Intervention Survey.** At the end of the three-week intervention, we assessed participants' PUMP and self-regulation ability using the same scales in a post-intervention survey. Additionally, for the PRC group, we employed the Godspeed scale to assess how participants perceived the chatbot regarding its intelligence and anthropomorphism (scale ranging from 1–7) [8].

**4.2.5 Debriefing Interview.** We conducted semi-structured interviews with each participant at the end of the intervention. Participants were asked about their overall experience with StayFocused, the scenarios where they usually started a focus session, what they liked or disliked about StayFocused, and how they felt about their smartphone use behaviors. For participants in PR and PRC groups, we also asked about their reactions to the reflective prompts and impressions of the chatbot. For contextualization, we encouraged them to refer to their focus logs in StayFocused during the interview. Each interview lasted 15 to 30 minutes.

**4.2.6 Collecting Post-Intervention Smartphone Use.** One week (7 days) after the intervention, we contacted each participant, asking them to log back into StayFocused so that their smartphone use data during this period would be uploaded to our database.

### 4.3 Data Collection and Analysis

Our study collected both quantitative and qualitative data from the app logs, smartphone use duration, surveys, and interviews. This mixed approach allowed us to understand how StayFocused affected participants' phone-checking behaviors during intended focus time, as well as the effects of the reflective prompts and chatbot support on their overall phone use habits from multiple perspectives.

**4.3.1 Survey Data.** We gathered participants' self-reported PUMP and self-regulation ability before and after the intervention stage in a survey. To examine these two metrics pre and post-intervention, we used paired *t*-tests.

**4.3.2 StayFocused Usage.** Participants' use of StayFocused included the number of focus sessions they started and completed, their planned and actual focus duration of each session, attempts to leave the focus sessions earlier, and records of returning to the session after attempting to leave. To understand whether and how these behaviors differed in each group, we used mixed-effects models to handle the unbalanced data with repeated measures and treated participant as a random effect; then we performed the Bonferroni test to examine specific group variations. This method has been widely used in previous HCI research on human behavior analysis in real-world settings [38, 51].

**4.3.3 Smartphone Usage.** StayFocused automatically gathered participants' daily smartphone use including screen duration and unlock frequency one week before, during, and one week after using the app (see Section 3.3 for implementation details). Due to the varied settings of different Android manufacturers, we faced technical difficulties in capturing two PR and two PRC participants' screen durations. As a result, these four participants were excluded from the phone usage analysis. Additionally, we noticed some outliers in the remaining participants' daily screen duration (e.g., extremely short such as 10 minutes, or extremely long such as 18 hours). Therefore, we conducted Interquartile Range (IQR) analysis for each participant during each study stage (e.g., a participant may have different phone use habits pre, during, and post intervention), and removed 70 (6.85%) outliers that are below the lower bound ( $Q1 - 1.5 IQR$ ) and above the upper bound ( $Q3 + 1.5 IQR$ ) from 1022 data points. After data cleaning, we first employed linear regression to examine the factors (e.g., age, gender, pre-intervention PUMP) that might be strongly related to individual participants' phone screen duration

before they used StayFocused, which helped us identify the covariates that needed to be included in the subsequent analysis. Next, we compared participants' smartphone use behavior across three groups and different stages (pre-intervention, intervention, and post-intervention) and groups using mixed-effects models, while treating individual participants as random effects.

**4.3.4 Responses to Reflective Prompts and Conversation Logs.** We collected participants' responses to the reflective prompts (PR group) and conversation logs with the chatbot (PRC group). To investigate whether the presence of a chatbot could encourage more active responses, we compared the two groups regarding the number of questions answered and the number of words in each answer using independent *t*-tests. We also performed sentiment analysis of participants' responses using TextBlob [72], a widely used sentiment analysis tool that has shown to be consistent and accurate [11]. By assigning a score ranging from -1 to 1 to each sentence of the response, TextBlob determines that a score close to -1, 0, and 1 indicates negative, neutral, and positive sentiment, respectively. Examples of positive and negative responses from our participants are provided in Section 5.3.1. We then employed independent *t*-tests to examine whether the participants' responses to the reflective prompts (where they attempted to leave the session and where they completed the session) exhibited different emotional patterns in the two groups. To further understand why participants want to leave the focus session earlier before it ends, we conducted a content analysis of their responses to the first reflective prompt on phone-checking intentions during focus sessions.

**4.3.5 Interview Data.** We audio-recorded all the interviews and transcribed them into text. Three researchers worked together to analyze the interview data following the steps in thematic analysis [12, 13]. First, we familiarized ourselves with all the data and independently coded the same four transcripts (36.36%) to generate an initial list of codes. We then iteratively discussed discrepancies and merged the codes with the same meanings during our regular meetings. Next, two researchers divided up the work to complete analyzing the remaining transcripts and cross-check their codes. Last, the three researchers collaboratively organized the codes into emerging themes. While the nature of the analysis is bottom-up, we paid particular attention to (1) what participants usually did for the focus session, (2) why they wanted to leave the focus session earlier before reaching their goals, (3) whether and how the StayFocused app helped them reduce compulsive smartphone use during the focus sessions, and (4) what the challenges regarding compulsive smartphone use that StayFocused couldn't help to address.

## 5 RESULTS

During the three-week deployment of StayFocused, we collected a series of quantitative data including 2461 focus session records (832 in PB, 833 in PR, and 796 in PRC groups), participants' smartphone usage (7-day prevention, 21-day intervention, and 7-day post-intervention stages), and their self-reported smartphone use behaviors. We also collected qualitative data including participants' responses to the reflective prompts (PR group), conversation logs with the chatbot (PRC group), and their subjective experience with StayFocused from the interviews. Here, we organize the findings into three sections on participants' focus session activities, their



overall smartphone use, and experience in using StayFocused to manage their smartphone use behaviors.

## 5.1 Focus Session Activities

This section describes how participants used StayFocused to plan and complete their goals by resisting smartphone distractions. We first provide an overview of their focus session completion and duration, and then delve into their attempts to leave the focus sessions earlier by checking their phones.

**5.1.1 Focus Session Completion.** All three groups achieved over 80% completion rate of their focus sessions on average (PB: 81.57%, PR: 82.79%, PRC: 82.96%), without significant difference across groups (See Table 2). When asked what types of activities participants normally engaged in during their focus session, most participants reported doing work- and study-related tasks, such as homework (PR-3, PR-12), self-learning (PR-12, PR-13, PRC-8), and reading (PB-1, PB-9, PB-11, PRC-1, PRC-3, PRC-10). Some participants leveraged the focus session to keep themselves concentrated during lectures or group meetings (PB-3, PR-3, PR-5, PRC-8). Occasionally, participants simply wanted to stay away from their smartphones and normally started a focus session while doing chores (PRC-10), taking a mental break (PB-1, PR-10, PRC-5), and doing exercise (PB-10, PB-11, PR-7). Participants in all three groups indicated that setting up a focus session on StayFocused enabled them to “submit a commitment,” which dedicated themselves to tasks at hand and stay away from smartphone distractions (PB-1, PB-10, PR-1, PR-4, PR-5, PR-7, PR-9, PR-11, PR-13, PRC-2, PRC-3, PRC-4, PRC-7). For example, PRC-3 said: “I was able to fully focus on an activity without being distracted by my phone” (PRC-3). Similarly, PR-9 mentioned that “(the focus timer) helps you focus on whatever you are doing” (PR-9).

**5.1.2 Planned and Actual Focus Duration.** For most focus sessions, participants often set 25 minutes as their goal of focus duration and tended to focus longer than they planned (see Table 2). There was no significant difference among the three groups regarding their planned and actual focus duration *per session*. However, from the mixed-effects model, participants in the PR and PRC groups focused longer *per day* than those in the PB group, with a marginal significance (PR:  $M = 143$ ,  $t = 1.705$ ,  $p = .089$ ; PRC:  $M = 152$ ,  $t = 1.731$ ,  $p = .086$ ). During the interviews, participants in the PR and PRC groups (PR-4, PR-5, PR-9, PR-13, PRC-4, PRC-7) were aware of their tendency to focus longer than planned: “I normally exceed the time I set in order to concentrate longer on vital activities” (PRC-4). PRC participants also mentioned that they were encouraged to focus longer by the chatbot because sometimes it would ask whether they had planned to focus longer in the future (PRC-7). Additionally, we found that for all the participants, their pre-intervention self-regulation scores played a significant role in their actual focus duration: those with higher self-regulation scores tended to focus longer *per day* than those with lower self-regulation scores ( $t = 4.175$ ,  $p < .001$ ). Note that participants’ pre-intervention self-regulation scores showed no significant difference across groups.

As the study progressed, we did not observe significant changes in participants’ focus duration *per session*, but we found participants’ focus duration *per day* significantly increased ( $t = 2.542$ ,  $p = .011$ ). Echoing this result, participants acknowledged that they had a hard

time focusing at the beginning of the study (PB-1, PB-10, PR-7, PR-9, PR-11, PR-13, PRC-2), but were able to overcome the distractions and got used to staying in the focus session later: “As time went on, I was able to stay focused, I was able to keep my phone away from me like, even beyond the normal time of the focus, I was able to do other things like reading books” (PR-7).

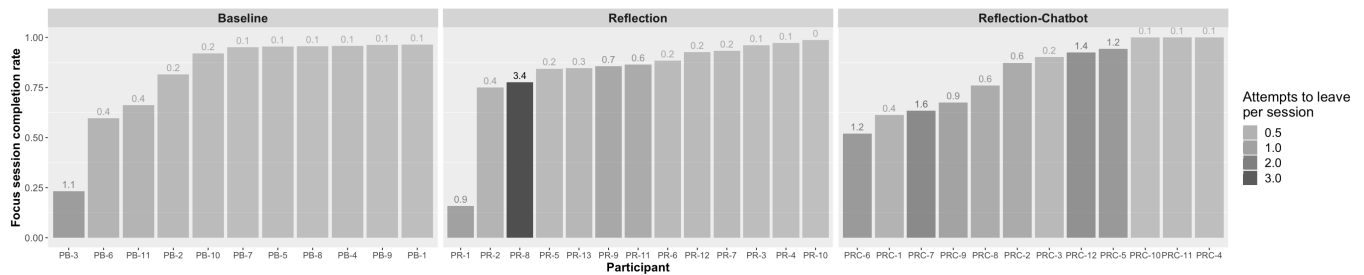
**5.1.3 Attempts to Leave the Focus Session Earlier.** During the focus sessions, participants in the PRC group attempted to leave the focus sessions more frequently compared to those in the PB ( $t = -4.625$ ,  $p < .001$ ) and PR groups ( $t = -3.373$ ,  $p = .002$ ), while no significant difference between participants in the PB and PR groups was observed. As shown in Table 2, PRC participants also had a significantly higher chance of returning to the focus sessions during their leaving attempts, compared to PB ( $t = -7.255$ ,  $p < .001$ ), and PR ( $t = -4.506$ ,  $p < .001$ ). Furthermore, although participants in the PB and PR groups did not differ regarding their average leaving attempts, those in the PR group had a higher chance of returning to the focus sessions than those in the PB group ( $t = -2.971$ ,  $p = .009$ ). This finding also emerged from our interviews, as both PR and PRC participants believed that the reflective prompts played an important role in “persuading” them to stay focused. For instance, PR-6 said “(it was because) the fact that if I try to leave the app, it keeps reminding me that I have a commitment to keep up with.” In Section 5.3, we provide a more detailed description of how the reflective prompts affected participants’ thoughts and practices during their attempts to leave the focus session. Besides, we found that for the PB group, those with a higher number of attempts to leave earlier tended to have a lower completion rate ( $cor = -.72$ ,  $p < .001$ ). Whereas for the PR and PRC groups, we did not observe such a correlation. Figure 4 illustrates individual participants’ completion rate while distinguishing their average number of attempts to leave the session earlier.

**5.1.4 Phone-Checking Intention.** We collected the reasons why participants wanted to check their phones during the focus sessions through their responses to the first reflective prompt (e.g., “Why do you want to check your phone right now?”), and further grouped them into five categories ( $n = 244$ ): (1) non-emergent urges to use the phone (25, 10.25%), including checking social media (PB-8, PB-10, PR-1, PR-6, PR-8, PRC-3, PRC-6, PRC-7, PRC-8), searching information (PR-2, PRC-3), and completing chores such as scanning QR code or make a bank transfer (PR-1, PRC-7); (2) sending or replying to calls and messages (83, 34.02%); (3) needs to check the phone as a transition between activities (50, 20.49%), such as relaxation after work (PR-1, PR-2, PRC-6, PRC-7, PRC-8, PRC-9); (4) using a productivity app (8, 3.28%), including alarm (PR-1, PR-2), digital notes (PRC-8), and calculator (PR-8, PRC-8); and (5) inherent difficulty in staying focused (36, 14.75%), due to tiredness (PR-2) and boredom (PRC-5, PRC-6, PRC-9). There were also instances that did not explicitly explain why a participant wanted to check their phones (42, 17.21%), suggesting that participants sometimes might not take the questions seriously or lacked awareness of their phone-checking intentions. Some of these instances are “Use it” (PR-2), “to end” (PR-13), “no reason” (PRC-5), “I don’t know why” (PRC-1), an emoji, or punctuation marks (PRC-8, PRC-12).

## 5.2 Smartphone Use Behavior

**Table 2: Participants’ focus behavior with StayFocused, including the number of focus sessions they started, session completion rate, planned focus duration, actual focus duration, the chance of attempting to leave the session earlier (this value is smaller than 1, because in some focus sessions, participants did not attempt to leave), and the chance of returning to the focus session while attempting to leave. The data are rounded and aggregated on a per-participant basis by mean.**

Group	Session started	Completion rate	Planned focus duration per session (min)	Actual focus duration per session (min)	Actual focus duration per day (min)	No. of attempts to leave earlier per session	Chance of returning to the focus session per attempt
PB ( $n = 11$ )	76	0.82	25	36	135	0.30	0.40
PR ( $n = 13$ )	64	0.83	25	44	143	0.40	0.57
PRC ( $n = 12$ )	66	0.83	28	47	152	0.65	0.75



**Figure 4: Participants’ focus session completion rate arranged in ascending order. The transparency of color represents each participant’s tendency to leave a focus session earlier before it ends. Darker colors indicate more attempts to leave the session (the data label above the bars indicates the average number of attempts made by each participant to leave a focus session earlier).**

**Table 3: Average and standard deviation of participants’ phone screen duration (mins) before (one week), during (three weeks), and after (one week) using the StayFocused app. Statistically significant changes ( $p < .05$ ) are marked with \*. Four participants’ data (2 in PR, and 2 in PRC) were excluded from the analysis due to data capture errors.**

Group	Pre-intervention	Intervention	Intervention vs. pre	Post-intervention	Post vs. intervention	Post vs. pre
PB ( $n = 11$ )	$M = 514, SD = 139$	$M = 423, SD = 123$	$M = -91 (17.70\%) *$	$M = 493, SD = 106$	$M = +70 (16.55\%) *$	$M = -21 (4.09\%)$
PR ( $n = 11$ )	$M = 595, SD = 201$	$M = 546, SD = 219$	$M = -49 (8.24\%)$	$M = 557, SD = 171$	$M = +11 (2.01\%)$	$M = -38 (6.39\%)$
PRC ( $n = 10$ )	$M = 571, SD = 227$	$M = 509, SD = 225$	$M = -62 (10.86\%)$	$M = 489, SD = 85$	$M = -20 (3.93\%)$	$M = -82 (14.36\%)$

Before using StayFocused, our participants spent 560 minutes ( $SD = 184$ ) per day on their smartphones without significant differences among the groups. This average duration translates to 9 hours and 20 minutes, which was about three hours longer than the participants’ self-estimated screen duration in the pre-study survey. On average, participants’ pre-intervention PUMP score is 2.39 out of 7, indicating a medium to high level of problematic phone use habit [55]. This measure did not significantly differ among the three groups. The linear regression model showed that participants with a lower PUMP score (more problematic phone use) indeed used their phones longer ( $t = -5.899, p < .001$ ).

Table 3 lists the screen duration of the three groups during each stage, along with their corresponding changes. Compared with the pre-intervention stage, all participants’ phone screen duration decreased during the intervention stage (PB:17.70%, PR: 8.24%; PRC: 10.86%), but according to the post-hoc test results, only the PB group exhibited a significant decrease ( $t = 4.02, p = .002$ ). Despite the general decreasing trend in screen duration during the intervention stage, we noticed some exceptions whose screen duration increased (PB-2, PB-4, PB-11, PR-5, PR-7, PR-9, PRC-8, PRC-10, PRC-12). Post-intervention, the screen duration relapsed for the PB and PR groups

(PB: 16.55%, PR: 2.01%), and the change was significant only within the PB group ( $t = -3.18, p = .05$ ); for the PRC group, on the contrary, participants’ screen duration further decreased (3.93%), although the change was not significant. Comparing the screen duration pre- and post-intervention, the PRC group exhibited the most substantial reduction (14.35%). Figure 5 illustrates the aggregated screen duration of all the participants as the study progressed. The screen unlock frequency was not included for analysis, as the data did not necessarily reflect participants’ smartphone use behavior. For example, we observed a slight increase in all the participants’ screen unlocks during the intervention stage without significant changes (PB = +5.5, PR = +3.5, PRC = +4.6), which was likely caused by the focus session-related interactions.

At the end of the three-week intervention, all participants’ post-intervention PUMP scores significantly increased, suggesting an improvement in their problematic phone use (PB:  $M = .78, t = 2.60, p = .025$ ; PR:  $M = .59, t = 2.08, p = .059$ ; PRC:  $M = 2.08, t = 4.07, p = .002$ ). Notably, the PRC group showed the most substantial reduction. Participants also reported perceiving less smartphone use during



**Figure 5: Changes in screen duration (mins) of the three groups throughout the pre-intervention, intervention, and post-intervention stages. As shown in the mixed-effects model, on average, PB participants’ phone screen duration significantly reduced from the pre-intervention to the intervention stage, but significantly relapsed post-intervention. For PR and PRC groups, their screen duration also decreased from the pre-intervention to the intervention stage, although the changes were not significant. In the post-intervention, the PR group had a slight relapse, while the PRC group further decreased their screen duration. However, we noticed the large individual variances in participants’ screen duration data, especially towards the end of the post-intervention stage, and elaborated on the possible reasons later in the Discussion.**

the intervention stage (PB-1, PB-11, PR-3, PR-9, PR-11, PRC-4, PRC-7, PRC-8, PRC-9). For example, PB-1 believed that they used their phone less in order to “*be concentrated*.” PR-3 found themselves able to stay away from their phones “*even without starting a focus session*,” and PRC-4 highlighted that “*the app instilled some level of discipline in me, reduced my phone addiction*.” However, we did not observe a significant change in participants’ self-regulation scores in all three groups.

### 5.3 The Effects of reflective prompts

StayFocused gathered a total of 1,602 response sets from the PR (788) and PRC (814) groups. A set refers to a series of responses that occurred when participants attempted to leave a focus session or when they completed a session; each set consists of one to four responses, depending on whether participants skipped any questions or not. Most of these response sets (1,358, 84.77%) were collected upon session completion (PR: 702; PRC: 656), and 244 sets (15.23%) were collected when participants attempted to leave the session earlier (PR: 86, PRC: 158). Within the latter sets, 228 of them are related to the case where participants eventually left the session (PR: 85, PRC: 143) and only 16 are related to the case where participants returned to the session during their attempts to leave (PR: 1; PRC: 15).

As shown in Table 4, it is noteworthy that there were 522 instances (not included in the response sets) where participants returned to the focus session during their attempts to leave the focus session without responding to the reflective prompts (PR: 173, PRC: 349). This suggested that among the total of 538 returns during the attempts to leave, 97.03% occurred when participants did not answer any prompts. For the PB group, we collected 66 returning records from 204 attempts to leave; although this group did not receive any reflective prompts, they still needed to confirm whether to leave or return before the focus session ended. On average, both PR and PRC participants spent less than 2 minutes responding to the reflective prompts.

**5.3.1 Response Characteristics.** In the session completion scenario, participants in the PR group answered slightly more questions than those in the PRC group (PR: 3.95, PRC: 3.52,  $t = 12.51$ ,  $p < .001$ ), while the PRC group’s responses included more words than the PR group (PR: 1.91, PRC: 2.85,  $t = 10.00$ ,  $p < .001$ ). In the scenario where participants attempted to leave the focus session, the two groups did not differ in the number of questions they answered (PR: 3.08, PRC: 3.14), but the PRC group responded with more words (PR: 2.53, PRC: 4.01,  $t = 4.78$ ,  $p < .001$ ).

Through sentiment analysis with TextBlob, we found that participants tended to react more positively when they completed a session than when they attempted to leave a session (completion: .26, leaving: .06,  $t = 16.58$ ,  $p < .001$ )<sup>2</sup>. During their attempts to leave the focus session, PRC participants showed slightly more positive sentiment in their responses than those in the PR group (PR: 0.03, PRC: 0.08,  $t = 2.65$ ,  $p = 0.009$ ), no matter whether they returned or not. Typical examples of positive responses included “*I was able to stay away from my mobile and do my chores, I felt all good and relaxed*.” (PRC-10), and “*It went smoothly*” (PR-7); and negative responses included “*I don’t want to have a plan for this*” (PRC-8) and “*boring, tired, anxious*” (PRC-4).

**5.3.2 Enabling Situated Reflection During Attempts to Leave the focus session.** As we reported in Table 2, participants in the PR and PRC groups had a higher chance of returning to the focus sessions when they attempted to leave, even though they did not necessarily answer the reflective prompts every time. This finding was corroborated during the interviews, where participants found themselves being prompted to think about their phone-checking intentions and their focus plans, which made them reconsider their next steps (PR-4, PR-6, PR-9, PR-11, PR-12, PRC-1, PRC-2, PRC-3, PRC-7, PRC-9). For instance, PR-12 said: “*they enable me to reflect on the sessions, like what I had done and what I am going to do*.” Some participants even explicitly mentioned how the prompts changed

<sup>2</sup>Sentiment scores close to -1, 0, and 1 indicate negative, neutral, and positive sentiment, respectively (see Section 4.3.4).

**Table 4: Participants' attempts to leave the focus session earlier and whether they returned to the session with or without responding to the reflective prompts (Note that the data presented in this table are aggregated based on each group, which slightly differs from the data in Table 2 that aggregates the data based on the individual participant.).**

Group	Total attempts	Returned with responses	Returned without responses	Left with responses	Left without response
PB ( $n = 11$ )	204	NA	66 (32.35%)	NA	138 (67.65%)
PR ( $n = 13$ )	305	1 (0.33%)	173 (56.72%)	85 (27.87%)	46 (15.08%)
PRC ( $n = 12$ )	508	15 (2.95%)	349 (68.70%)	143 (28.15%)	1 (0.20%)

their mind: “by looking at this question I changed my mind, and just went back to the focus session. (PR-11).”

PR-6 and PR-9 further elaborated on their experience with the reflective prompts, acknowledging that they sometimes felt guilty about not staying up with their commitment: “when I try to leave, the questions it asked make me feel bad like I am not keeping to something I planned or committed to doing” (PR-6). PRC participants also noted the similar feelings: “The advantages of using chatbot are that I was able to reflect about my focus session and make better resolution” (PRC-3).

**5.3.3 Providing Encouragement Upon Session Completion.** Compared with the response sets during attempts to leave the focus sessions, StayFocused gathered more response sets upon session completion. In this scenario, participants engaged with the prompts more actively and positively, receiving sustained support and encouragement to improve their focus habits. For example, PR-4 highlighted a connection built through the completion of each session, stating: “after completing the session, I felt connected to the app and it really felt awesome.” Additionally, the reflective prompts such as “How did the focus session go?” and “Does the complete session bring you closer to your goal today?” were found to be encouraging. PR-6 noted these prompts helped them build up a record of accomplishment: “Once I finish, the questions it asked make me feel relief like I am doing something good to bring myself together.”

## 5.4 The Effects of Chatbot Support

While the previous section already covers some of the PRC group's experience in responding to the reflective prompts, this section specifically focused on this group's unique experience with chatbot support. Overall, participants in the PRC group perceived the chatbot in StayFocused, Alan, as moderately intelligent ( $M = 2.50$ ,  $SD = 1.33$ ) and moderately anthropomorphic ( $M = 2.62$ ,  $SD = 0.83$ ). Among these participants, three perceived the chatbot as highly intelligent (PRC-1, PRC-3, PRC-6) and three perceived it as highly anthropomorphic (PRC-1, PRC-6, PRC-9), while another three perceived it with both low intelligence and low anthropomorphism (PRC-2, PRC-4, PRC-8). While participants' perceived intelligence and anthropomorphism did not seem to affect their focus behavior (as shown in Figure 4, PRC-2 and PRC-4 both achieved over 80% focus session completion rate), from the interviews, we found these perceptions greatly influenced individuals' engagement with the chatbot and their willingness to continue using StayFocused.

**5.4.1 Increasing Accountability.** Those who perceived the chatbot as moderately to highly anthropomorphic and intelligent expressed a feeling of being “observed,” as PRC-9 described: “when I leave a focus session, it sometimes feels like the bot, kinda yelling lol, like

why did you check your phone.” PRC-1 also recalled a similar feeling: “as for feeling observed, sometimes I would tell the chatbot what I did during the study session, so it served as a reminder and recorder of my behavior.” In this case, the participant found that sharing their daily activities with the chatbot helped them stay accountable for their study and work. In addition, PRC-7 mentioned that they learned several informative coping strategies from the chatbot to overcome distractions: “it has taught me how to set aside things that are distracting me at that point, aside from my phone.” Some examples of the strategies such as “plan your focus sessions for a time when you are more alert” and “maybe you can set a limit for yourself? For example, checking your phone every 10 minutes or so,” helped the participant to stay more productive.

**5.4.2 Receiving Emotional Support.** The above participants also expressed a feeling of being understood and emotionally supported. For example, there were times when PRC-1 needed to leave the session earlier due to tiredness; instead of keeping them in the focus session, the chatbot said: “it is great that you have persisted for a long time, and it is understandable that you feel tired. Maybe you can take a break and then come back to continue your work later.” PRC-6 and PRC-9 appreciated the encouraging words from the chatbot, which further strengthened their sense of accomplishment: “when I completed a session, the chatbot was happy for me and proud for achieving my goal of focusing for the time I set. And that encouraged me” (PRC-9). PRC-7 added that the chatbot was supportive, because “it's more like I am talking to a human being who wants to see me grow and succeed, the motivation keeps me going.” Oftentimes, the conversations with the chatbot could go beyond the focus session. PRC-1, PRC-5, PRC-7, and PRC-10 frequently shared their work and study-related stress (e.g., preparing for job interviews and exams) with the chatbot, which helped them “refresh from tiredness.” For example, when PRC-5 told the chatbot about their upcoming job interviews, the chatbot replied: “I am sure you will do great!”

**5.4.3 Engaging in Chit-Chat.** At times, participants chose not to adhere to the prompts presented by the chatbot and instead engaged in “testing its responses and conversational abilities.” PRC-5, for instance, asked questions such as “How many people you are helping with?” or responded with exclamatory words such as “Boooooommm” when the chatbot checked how their focus session went. In response to these off-topic comments, the chatbot was able to switch to a more lighthearted tone by saying “Haha; Great!” There were also cases where participants talked to the chatbot in non-English languages. For example, PRC-1 felt talking with the chatbot in English reminded them of work due to English being the official language at school. But one day when they “decided to respond to the prompts in Chinese (their native language) and see what would

happen” the chatbot understood and responded appropriately in Chinese, which reignited their motivation to use the app.

**5.4.4 Frowning Upon the “Stubbornness”.** Participants who perceived the chatbot as low-intelligent pointed out its limited conversational ability, such as appearing to be disconnected from the ongoing context and being too “stubborn” by always trying to find out why participants checked their phones. PRC-8, who held a negative impression of the chatbot, said “*it (has) very very strong assumptions. That is you should not touch your phone during your work.*” This lack of flexibility and contextual understanding made PRC-8 feel that talking to the chatbot was “an additional task.” Similarly, PRC-2, despite feeling the app was helpful, also felt that the chatbot asked too many questions without understanding their situation. From the conversation logs, we found several restless or even irritated responses when asked about one’s phone-checking intention: “*Is that all for the session?*” (PRC-12), “*I need my phone now!*” (PRC-8), “*I just wanna to leave-leave it away-*” (PRC-1).

Even those who generally liked the chatbot found it was not perfect in handling their negative emotions. As PRC-1 recalled some frustrating moments when they were really tired from work, the chatbot still kept asking them to stay focused: “*It’s not about the vocabulary or politeness. I mean, the chatbot is actually very polite, but it’s more about it being too stubborn if I can say that. Sometimes I felt all it cared about was that focus session. It would assure me that I would feel better later if I stuck to it.*”

**5.4.5 Faded Novelty Effects.** Participants also mentioned that the novelty effects of the chatbot faded away over time (PRC-1, PRC-2, PRC-5, PRC-6, PRC-8), as PRC-1 described: “*the first week (of interacting with the chatbot) was really interesting. But the longer time I used it, the less patience I have*” (PRC-1). Likewise, PRC-5 was excited and curious about the chatbot at the beginning of the study—they started four to six focus sessions daily and engaged in the conversations with the chatbot (e.g., greeting the chatbot upon completing a focus session by saying: “*Hi Alan*” and asking the chatbot about its activities: “*So what did u do today. Let’s share*”). But as the study progressed, their interest gradually diminished, which manifested in their reduced daily focus sessions (two to three) and passive responses such as “*Good*,” “*yo*,” or just a single punctuation mark. Some participants also complained about the repetitiveness of the prompts although we randomized and paraphrased them every time (PRC-2, PRC-6, PRC-8): “*Well, sometimes it is kind of frustrating because you see the same questions again and you wanna leave as soon as possible*” (PRC-6).

## 6 DISCUSSIONS

### 6.1 Reflecting on the Focus Session Activities and Smartphone Use

Our data did not show a significant difference in participants’ focus session completion rate among the three groups. Combining the findings around participants’ focus duration, their phone-checking behaviors during the focus sessions, and overall smartphone use, our interpretation of the statistical analysis results involved three aspects. First, our participants were all college students who were motivated to improve their smartphone use habits; despite self-identified problematic smartphone use, they were able to set reasonable focus goals and follow the plan most of the time, which is

evidenced by the average focus session completion rate exceeding 80%. Second, as shown in Figure 5, the smartphone use patterns of the PR and PRC groups appear to be less consistent. We suspect the presence of some “extreme” users in these two groups, such as those who were impatient with the reflective prompts (e.g., PR-1) or held negative impressions of the chatbot (e.g., PRC-8), exhibited different behaviors than others within the same group. Relatedly, during the focus sessions, the PR and PRC groups made more attempts to leave earlier than the PB group (see Table 2), which leads us to question if we overlooked any important factors (e.g., personality, lifestyle, cultural background) other than the PUMP and self-regulation scales in the group assignment or the option to explain one’s phone-checking intentions made these participants perceive the focus sessions more flexible. Although the reflective prompts and chatbot support did not necessarily increase the session completion rate, they could have potentially helped those who were more restless during the focus sessions (e.g., PR-8, PR-9, PR-11, PRC-5, PRC-6, PRC-7, PRC-12; see Figure 4) to improve their ability to focus. This can be inferred by the finding that PB participants who attempted to leave the focus session more often had a lower completion rate, but this correlation was not observed in the PR and PRC groups. Third, our data suggested that the reflective prompts and chatbot support potentially encouraged longer focus duration, which was also noticed by the PR and PRC participants themselves. These participants expressed a feeling of being “observed,” making them more self-conscious when committing to their focus goals. However, it is important to note that the marginal  $p$ -value warrants further investigation to demonstrate the replicability and generalizability of this relationship.

Regarding smartphone use behaviors, our results showed that the PB group reduced their screen duration more than the PR and PRC groups during the intervention stage. One possible reason was that the PR and PRC groups, having focused for longer durations during their focus sessions, subsequently spent more time on their phones later on to compensate for the “lost time.” This phenomenon aligns with prior research on ego depletion, which showed that prolonged periods of productivity can lead people to perceive a deficit in their leisure and entertainment time [27, 43]. However, it is worth mentioning that during the post-intervention stage, the PB group experienced a significant relapse in their screen duration reduction, indicating that solely relying on setting focus goals to avoid smartphone use is not a sustainable solution. On the contrary, the reduction in screen duration within the PR and PRC groups remained more stable; particularly, the PRC group further decreased their screen duration after withdrawing StayFocused (see Section 5.2). Although this further reduction was not statistically significant, the PUMP score of the PRC group showed the most substantial increase after the intervention, suggesting the potential of the chatbot to facilitate more sustainable change. Additionally, the non-significance result could be due to the contrasting trend in the screen duration of two PRC participants (PRC-4, PRC-8), which drastically increased from the intervention to the post-intervention stage. This within-group variation also highlighted the need to consider individual preferences for chatbot-based persuasion to avoid counterproductive consequences.

It is important to note that not all smartphone usage is driven by compulsive behaviors [36, 38, 46], as our participants occasionally received urgent calls or needed to use specific applications on their phones for their study and work. Therefore, the focus session completion rate and the amount of screen duration reduction, while being helpful references, may not be the only and the best evaluation standard to assess one’s problematic smartphone use. Rather than solely looking into the numeric measures [39, 44], it is more meaningful to examine how the design elements in a smartphone non-use application enable individuals to set and execute focus plans, navigate distractions, reflect on phone-checking intentions, and successfully complete each session, etc. We believe this work contributes to such an understanding.

## 6.2 How Can reflective prompts + Chatbot Support Help Reduce Compulsive Smartphone Use?

**6.2.1 The “Silent Nudge”.** When the PR and PRC participants attempted to leave a focus session, the PB, PR, and PRC groups had 0.4, 0.57, and 0.75 likelihood to return, with significant differences (see Table 2), which demonstrated the effectiveness of the reflective prompts and the chatbot support. Interestingly, the majority of PR and PRC groups’ returns (about 97%) occurred without any responses to the reflective prompts; and when participants did respond to the prompts, it was often an indication that they were determined to leave (see Section 5.3). We suspect that if participants checked their phones simply out of habit without a specific reason, prompts such as “*why did you want to check your phone right now?*” or “*Why do you want to leave the focus session,*” could have prompted them to situationally shift their mindset back to focus. In this case, the mere presence of reflective prompts served as a “silent nudge” to reduce participants’ unconscious phone-checking behaviors, which indicated that situated reflection may not require explicit articulation or elaborate explanation of one’s intention. This finding echoes previous research on mindful smartphone use, which suggested that a simple prompting question could enable meaningful reflection, while people do not necessarily need to specify their behavior rationales [71]. Likewise, the goal-directed priming effects showed that external cues related to one’s goals can activate behavior change without conscious intentions [60]. In particular, we found that the effects of this silent nudge were more pronounced in the PRC group. This was likely due to the chatbot’s presence as a humanlike companion that increased participants’ accountability. Similar findings were shown in prior work that leveraged chatbot as a reflection companion to increase individuals’ awareness of their daily activities [40–42].

However, if participants had a strong urge to check the phones or genuinely needed them, they often proceeded to leave even after responding to the prompts. In this case, the reflective prompts and chatbot support were not sufficient to mitigate their phone-checking intention. While not as previous studies envisioned that articulating one’s intention could lead them to reconsider immediate actions [49, 52], our study showed that participants took the opportunity to justify rather than change their behaviors. We believe this justification process was still meaningful, as participants mentioned during the interviews, responding to the prompts made them realize

how much they are dependent on their phones, leading to a sense of guilt. This result aligned with prior work, which demonstrated that probing one’s behavior reasoning can be thought-provoking and deepen self-awareness [33, 49]. Interestingly, as Table 4 shows, the PRC group was more likely to provide their justification than the PR group when they eventually decided to leave, which could be due to their increased accountability affected by the chatbot.

**6.2.2 More Encouragement, Less Questioning.** Our results showed that PR and PRC participants engaged with the prompts upon focus session completion more often and more positively than during the attempts to leave a session earlier. This post-session engagement not only enabled self-reflection on the particular session but also extended the reflection to participants’ overall smartphone use habits and future plans (e.g., by asking “*how did you feel without smartphone distraction?*”, “*what is your next step?*”). Responding to the reflective prompts could have fostered a sense of accomplishment among participants, serving as positive reinforcement to acknowledge their efforts [19, 38]. When the prompts were incorporated into a natural conversation with the chatbot, they further signified participants’ progress by sending additional feedback and encouragement. For example, PRC-7 and PRC-9 felt that the chatbot was happy and proud of their achievements, which could forge a positive loop for them to better regulate their smartphone use. Aligning with our results, prior work also found that a chatbot equipped with emotional intelligence (e.g., empathetic and social skills) can enrich people’s interaction experience and encourage them to take intended actions [26, 63].

On the other hand, some participants found the reflective prompts annoying and lacked flexibility. In certain situations, smartphone usage may be necessary rather than compulsive, but the chatbot did not understand such contextual nuances and kept emphasizing the importance of staying focused. These negative impressions hindered our intended design outcome of enabling situated reflection and even resulted in counterproductive effects, leading some participants to use their phones more frequently than before. In constructing the chatbot persona using GPT-3, the major challenge we encountered was to ensure that it could cover most of the reflective prompts provided in the reflection version but also be an “*open-minded and empathetic*” personal companion. Learning from prior work prompting LLMs [76], we have added several rules in designing our chatbot (e.g., “*I never repeat myself*”, “*I do not judge how people spend their time but listen attentively to what they say*”). However, these rules did not take effect every time due to the uncertainty of prompting LLMs [78]. To improve the conversation experience, we should revisit our prompt design and carefully consider when to ask which questions in different situations. This may require the collection of more conversation samples that can be used to fine-tune the LLMs rather than simply relying on prompt engineering.

**6.2.3 Diversifying the Prompts.** In addition to the “*stubbornness and judgments*,” some participants found the prompts repetitive, despite our efforts to paraphrase and randomize them. The repetitiveness could contribute to the diminished novelty effects of the chatbot, speeding up the loss of patience and interest. Such faded novelty effects were also commonly observed in applications involving human-AI interaction: people often perceive the agent as

intelligent and empathic at the beginning but soon realized that it was just a machine following pre-defined patterns [23, 32]. To sustain individuals' interests and motivation, we could diversify the conversations based on individual progress instead of simply randomizing all the questions across the study. According to the stage theories of health behavior [77], individuals have different mental models during different stages towards behavior change and thus the support they need also varies. For example, when a person first starts to overcome smartphone addiction, the prompts can center around their phone-checking intentions to foster awareness. As they make progress, the prompts can then be extended to their daily plans and suggestions for further improvement. Individuals' data such as their completed focus sessions and screen duration, can also be included in the prompts to create a more personalized experience. This could establish a deeper connection between people and the chatbot [42], resulting in strengthened accountability and commitment. Additionally, leveraging the large knowledge base of LLMs, the chatbot can act beyond a "reflection companion." As Akinci et al. demonstrated in their work, productivity tools powered by generative AI can aid in task engagement and mitigate procrastination by supporting the continuation of interrupted work [7]. In this light, StayFocused could send people prompts not just about self-reflection, but also advice on task management.

### 6.3 Study Limitations and Future Work

First, our smartphone use data (screen duration) was based on whether the screen was on or off, which may be affected by each phone's auto-lock settings. Therefore, the screen duration data may not perfectly reflect participants' actual time spent on their phones. Second, our chatbot design was not perfect due to the repetitive prompts and a lack of contextual understanding as mentioned above, which can limit the findings regarding the chatbot's potential. Third, despite our efforts to recruit participants from diverse regions, those who completed the study were predominantly from the U.S. Thus, our findings may not be generalized to other cultures. It is noteworthy that out of the 50 participants we initially recruited, 13 dropped out during the study, highlighting the inherent challenges in "persuading" people to reduce compulsive smartphone use.

As the first step to exploring the opportunities for chatbots as a reflective agent to reduce unintended behaviors, we designed and developed StayFocused in the context of helping college students manage compulsive smartphone use during intended focus time. We examined how the reflective prompts and chatbot support played parts in enhancing individuals' situated reflection on their phone-checking intentions, along with rich insights into their experiences. Going forward, the chatbot's conversation ability can be improved with more sophisticated prompting strategies and diversified reflective prompts, which can potentially make StayFocused more effective in conveying the persuasive messages. Besides, it can be interesting to integrate other interaction modalities such as speech input to avoid direct screen contact and interaction burden [49, 51], which could also enable more in-depth reflection by allowing individuals to elaborate on their phone-checking intentions without manual input. The lessons learned from this work can also be extended to other contexts, where reflection plays important roles in promoting behavior change, such as consistent exercise [50] and healthy eating habits [10].

### 6.4 Conclusion

In this work, we designed and built StayFocused, a mobile app that guides people to plan and complete focus sessions for *not* interacting with their phones. To increase individuals' awareness of their smartphone use and reduce this behavior when it is not necessary, we incorporated reflective prompts probing people to reflect on their phone-checking intentions *in situ* and investigated whether using a chatbot to deliver the prompts can further enhance their reflection. In a stage-based between-subjects study, we recruited 36 college students and assigned them to three groups: PB, PR, and PRC, who were asked to use the three versions of StayFocused: baseline, reflection, and reflection-chatbot, respectively. Although the three groups did not differ in their focus session completion rate, the PR and PRC participants focused longer and exhibited a higher chance of returning to the session while attempting to leave. By qualitatively analyzing participants' responses to the reflective prompts and their study experience during the debriefing interviews, we found that the reflective prompts effectively enabled participants' reflection on their phone-checking behaviors, and the chatbot helped increase their accountabilities and provided emotional support. On the other hand, we identified several limitations and challenges in the chatbot design including the "Stubbornness" and faded novelty effects. Reflecting upon the findings, we discussed the nuances associated with individuals' preferences and the resulting observations. We also discussed how to improve the current designs of StayFocused to better encourage healthier use of smartphones. The insights gained from this study hold the promise to guide the development of interventions to mitigate other unintended behaviors beyond compulsive smartphone use.

### ACKNOWLEDGMENTS

We thank our participants for their time and interest in this study. We also thank Trung Hai Le for helping develop the baseline and reflection versions of StayFocused. This research was supported by City University of Hong Kong (# 9610597).

### REFERENCES

- [1] Yuna Ahn, Yilin Zhang, Yujin Park, and Joonhwan Lee. 2020. A Chatbot Solution to Chat App Problems: Envisioning a Chatbot Counseling System for Teenage Victims of Online Sexual Exploitation. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (2020). <https://api.semanticscholar.org/CorpusID:218483168>
- [2] Tuncay Akinci. 2021. Determination of Predictive Relationships Between Problematic Smartphone use, Self-Regulation, Academic Procrastination and Academic Stress Through Modelling. *International Journal of Progressive Education* 17, 1 (2021), 35–53. <https://eric.ed.gov/?id=EJ1286359>
- [3] Suliman S. Aljomaa, Mohammad F. AlQudah, Ismael S. Albarsan, Salaheldin F. Bakhiet, and Adel S. Abduljabbar. 2016. Smartphone addiction among university students in the light of some variables. *Computers in Human Behavior* 61 (2016), 155–164. <https://doi.org/10.1016/j.chb.2016.03.041>
- [4] Mohamed Basel Almourad, Amen Alrobai, Tiffany Skinner, Mohammed Hussain, and Raian Ali. 2021. Digital wellbeing tools through users lens. *Technology in Society* 67 (2021), 101778. <https://doi.org/10.1016/j.techsoc.2021.101778>
- [5] Alice Alphonse, Kezia Stewart, Jamie Brown, Olga Perski, et al. 2022. Exploring Users' Experiences With a Quick-Response Chatbot Within a Popular Smoking Cessation Smartphone App: Semistructured Interview Study. *JMIR Formative Research* 6, 7 (2022), e36869. <https://doi.org/10.2196/36869>
- [6] Android Digital Wellbeing. 2023. <https://www.android.com/digital-wellbeing/>.
- [7] Riku Arakawa, Hiromu Yakura, and Masataka Goto. 2023. CatAlyst: Domain-Extensible Intervention for Preventing Task Procrastination Using Large Generative Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19. <https://doi.org/10.1145/3544548.3581133>

- [8] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1 (2009), 71–81. <https://doi.org/10.1007/s12369-008-0001-3>
- [9] Timothy Bickmore and Toni Giorgino. 2006. Health dialog systems for patients and consumers. *Journal of biomedical informatics* 39, 5 (2006), 556–571. <https://doi.org/10.1016/j.jbi.2005.12.004>
- [10] Johnna Blair, Yuhuan Luo, Ning F Ma, Sooyeon Lee, and Eun Kyoung Choe. 2018. OneNote Meal: A photo-based diary study for reflective meal tracking. In *AMIA Annual Symposium Proceedings*, Vol. 2018. American Medical Informatics Association, 252. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6371351/>
- [11] Venkateswarlu Bonta. 2019. A comprehensive study on lexicon based approaches for sentiment analysis. *Asian Journal of Computer Science and Technology* 8, S2 (2019), 1–6.
- [12] Robert Bowman, Camille Nadal, Kellie Morrissey, Anja Thieme, and Gavin Doherty. 2023. Using Thematic Analysis in Healthcare HCI at CHI: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18. Using Thematic Analysis in Healthcare HCI at CHI: A Scoping Review
- [13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [14] Nancy A Cheever, Larry D Rosen, L Mark Carrier, and Amber Chavez. 2014. Out of sight is not out of mind: The impact of restricting wireless mobile device use on anxiety levels among low, moderate and high users. *Computers in Human Behavior* 37 (2014), 290–297. <https://doi.org/10.1016/j.chb.2014.05.002>
- [15] Chongyang Chen, Kem ZK Zhang, Xiang Gong, Sesia J Zhao, Matthew KO Lee, and Liang Liang. 2017. Understanding compulsive smartphone use: An empirical test of a flow-based model. *International Journal of Information Management* 37, 5 (2017), 438–454. <https://doi.org/10.1016/j.ijinfomgt.2017.04.009>
- [16] Pin-Chieh Chen, Min-Wei Hung, Hsueh-Sung Lu, Chien Wen Yuan, Nanyi Bi, Wan-Chen Lee, Ming-Chyi Huang, and Chuang-Wen You. 2022. This App is not for Me: Using Mobile and Wearable Technologies to Improve Adolescents' Smartphone Addiction through the Sharing of Personal Data with Parents. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15. <https://doi.org/10.1145/3491102.3517478>
- [17] Hyunsung Cho, DaEun Choi, Donghui Kim, Wan Ju Kang, Eun Kyoung Choe, and Sung-Ju Lee. 2021. Reflect, not regret: Understanding regretful smartphone use with app feature-level analysis. *Proceedings of the ACM on human-computer interaction* 5, CSCW2 (2021), 1–36. <https://doi.org/10.1145/3479600>
- [18] Francesco Cirillo. 2013. *The Pomodoro technique: Do more and have fun with time management*. FC Garage.
- [19] 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 2008 CHI Conference on Human Factors in Computing Systems*. 1797–1806. <https://doi.org/10.1145/1357054.1357335>
- [20] Éilish Duke and Christian Montag. 2017. Smartphone addiction, daily interruptions and self-reported productivity. *Addictive behaviors reports* 6 (2017), 90–95. <https://doi.org/10.1016/j.abrep.2017.07.002>
- [21] Jon D Elhai, Haibo Yang, and Christian Montag. 2020. Fear of missing out (FOMO): overview, theoretical underpinnings, and literature review on relations with severity of negative affectivity and problematic technology use. *Brazilian Journal of Psychiatry* 43 (2020), 203–209. <https://doi.org/10.1590/1516-4446-2020-0870>
- [22] Forest. 2023. <https://www.forestapp.cc/>.
- [23] Luke K Fryer, Mary Ainley, Andrew Thompson, Aaron Gibson, and Zelinda Sherlock. 2017. Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior* 75 (2017), 461–468. <https://doi.org/10.1016/j.chb.2017.05.045>
- [24] Yaoguo Geng, Jingjing Gu, Jing Wang, and Ruiping Zhang. 2021. Smartphone addiction and depression, anxiety: The role of bedtime procrastination and self-control. *Journal of affective disorders* 293 (2021), 415–421. <https://doi.org/10.1016/j.jad.2021.06.062>
- [25] Eun Go and S. Shyam Sundar. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* 97 (2019), 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- [26] Ted Grover, Kael Rowan, Jina Suh, Daniel McDuff, and Mary Czerwinski. 2020. Design and Evaluation of Intelligent Agent Prototypes for Assistance with Focus and Productivity at Work. In *Proceedings of the 25th International Conference on Intelligent User Interfaces (Cagliari, Italy) (IUI '20)*. Association for Computing Machinery, New York, NY, USA, 390–400. <https://doi.org/10.1145/3377325.3377507>
- [27] Martin S Hagger, Chantelle Wood, Chris Stiff, and Nikos LD Chatzisarantis. 2010. Ego depletion and the strength model of self-control: a meta-analysis. *Psychological bulletin* 136, 4 (2010), 495. <https://doi.org/10.1037/a0019486>
- [28] Xu Han, Michelle X. Zhou, Matthew Turner, and Tom Yeh. 2021. Designing Effective Interview Chatbots: Automatic Chatbot Profiling and Design Suggestion Generation for Chatbot Debugging. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021). <https://api.semanticscholar.org/CorpusID:233209707>
- [29] Christina N. Harrington and Lisa Egede. 2023. Trust, Comfort and Reliability: Understanding Black Older Adults' Perceptions of Chatbot Design for Health Information Seeking. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023). <https://api.semanticscholar.org/CorpusID:258217665>
- [30] Severin Haug, Raquel Paz Castro, Min Kwon, Andreas Filler, Tobias Kowatsch, and Michael P Schaub. 2015. Smartphone use and smartphone addiction among young people in Switzerland. *Journal of behavioral addictions* 4, 4 (2015), 299–307. <https://doi.org/10.1556/2006.4.2015.037>
- [31] 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* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 4746–4757. <https://doi.org/10.1145/2858036.2858403>
- [32] Weijiao Huang, Khe Foon Hew, and Luke K Fryer. 2022. Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning* 38, 1 (2022), 237–257. <https://doi.org/10.1111/jcal.12610>
- [33] Jinkyu Jang and Jinwoo Kim. 2020. Healthier life with digital companions: Effects of reflection-level and statement-type of messages on behavior change via a perceived companion. *International Journal of Human-Computer Interaction* 36, 2 (2020), 172–189. <https://doi.org/10.1080/10447318.2019.1615722>
- [34] Wenning Jiang, Jin Luo, Hannan Guan, Feng Jiang, and Yi-Lang Tang. 2022. Problematic mobile phone use and life satisfaction among university students during the COVID-19 pandemic in Shanghai, China. *Frontiers in public health* 9 (2022), 805529. <https://doi.org/10.3389/fpubh.2021.805529>
- [35] Chien-Hao Kao, Chih-Chieh Chen, and Yu-Tza Tsai. 2019. Model of multi-turn dialogue in emotional chatbot. In *2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)*. IEEE, 1–5. <https://doi.org/10.1109/BIGCOMP.2017.7881752>
- [36] Inyeop Kim, Gyuwon Jung, Hayoung Jung, Minsam Ko, and Uichin Lee. 2017. Let's Focus: Location-Based Intervention Tool to Mitigate Phone Use in College Classrooms. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (Maui, Hawaii) (UbiComp '17)*. Association for Computing Machinery, New York, NY, USA, 101–104. <https://doi.org/10.1145/3123024.3123165>
- [37] 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* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300927>
- [38] 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* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 272–283. <https://doi.org/10.1145/2858036.2858428>
- [39] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. 2011. How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3063–3072. <https://doi.org/10.1145/1978942.1979396>
- [40] Ahmet Baki Kocaballi, Shlomo Berkovsky, Juan C Quiroz, Liliana Laranjo, Huong Ly Tong, Dana Rezazadegan, Agustina Briatore, and Enrico Coiera. 2019. The personalization of conversational agents in health care: systematic review. *Journal of medical Internet research* 21, 11 (2019), e15360. <https://www.jmir.org/2019/11/e15360/>
- [41] Rafal Kocielnik, Daniel Avrahami, Jennifer Marlow, Di Lu, and Gary Hsieh. 2018. Designing for workplace reflection: a chat and voice-based conversational agent. In *Proceedings of the 2018 designing interactive systems conference*. 881–894. <https://doi.org/10.1145/3196709.3196784>
- [42] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection Companion: A Conversational System for Engaging Users in Reflection on Physical Activity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2, Article 70 (jul 2018), 26 pages. <https://doi.org/10.1145/3214273>
- [43] Geza Kovacs, Drew Mylander Gregory, Zilin Ma, Zhengxuan Wu, Golrokh Emami, Jacob Ray, and Michael S Bernstein. 2019. Conservation of procrastination: Do productivity interventions save time or just redistribute it?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12. <https://doi.org/10.1145/3290605.3300560>
- [44] Simone Lanette, Phoebe K Chua, Gillian Hayes, and Melissa Mazmanian. 2018. How much is 'too much'? The role of a smartphone addiction narrative in individuals' experience of use. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–22. <https://doi.org/10.1145/3274370>
- [45] Young-Jun Lee, Chae-Gyun Lim, Yunsu Choi, Ji-Hui Lm, and Ho-Jin Choi. 2022. PERSONACHATGEN: Generating personalized dialogues using GPT-3. In *Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge*. 29–48. <https://aclanthology.org/2022.ccgpk-1.4>



- [46] Yu-Kang Lee, Chun-Tuan Chang, You Lin, and Zhao-Hong Cheng. 2014. The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. *Computers in Human Behavior* 31 (2014), 373–383. <https://doi.org/10.1016/j.chb.2013.10.047>
- [47] Zhuoyang Li, Minhui Liang, Hai Trung Le, Ray Lc, and Yuhan Luo. 2023. Exploring Design Opportunities for Reflective Conversational Agents to Reduce Compulsive Smartphone Use. In *Proceedings of the 5th International Conference on Conversational User Interfaces*. 1–6.
- [48] Yu-Hsuan Lin, Yu-Cheng Lin, Sheng-Hsuan Lin, Yang-Han Lee, Po-Hsien Lin, CL Chiang, Li-Ren Chang, CCH Yang, and TBJ Kuo. 2017. To use or not to use? Compulsive behavior and its role in smartphone addiction. *Translational Psychiatry* 7, 2 (2017), e1030–e1030. <https://doi.org/10.1038/tp.2017.1>
- [49] Yuhan Luo, Young-Ho Kim, Bongshin Lee, Naeemul Hassan, and Eun Kyoung Choe. 2021. FoodScrap: Promoting Rich Data Capture and Reflective Food Journaling Through Speech Input.. In *Proceedings of the 2021 Conference on Designing Interactive System*. ACM. <https://doi.org/10.1145/3461778.3462074>
- [50] Yuhan Luo, Bongshin Lee, and Eun Kyoung Choe. 2020. TandemTrack: shaping consistent exercise experience by complementing a mobile app with a smart speaker. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 1–13. <https://doi.org/10.1145/3313831.3376616>
- [51] Yuhan Luo, Bongshin Lee, Young-Ho Kim, Eun Kyoung Choe, et al. 2022. Note-Wordy: Investigating Touch and Speech Input on Smartphones for Personal Data Capture. *proc. of the ACM on Human-Computer Interaction, Interactive Surfaces and Spaces (ISS), (To appear)*. <https://doi.org/10.1145/3567734>
- [52] Yuhan Luo, Bongshin Lee, Donghee Yvette Wohn, Amanda L Rebar, David E Conroy, and Eun Kyoung Choe. 2018. Time for break: Understanding information workers' sedentary behavior through a break prompting system. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14. <https://doi.org/10.1145/3173574.3173701>
- [53] Ulrik Lyngs, Kai Lukoff, Petr Slovak, William Seymour, Helena Webb, Marina Jirotko, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2020. 'I Just Want to Hack Myself to Not Get Distracted': Evaluating Design Interventions for Self-Control on Facebook. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3313831.3376672>
- [54] Jocelyne Matar Boumoleh and Doris Jaalouk. 2017. Depression, anxiety, and smartphone addiction in university students- A cross sectional study. *PLOS ONE* 12 (08 2017). <https://doi.org/10.1371/journal.pone.0182239>
- [55] Lisa J Merlo, Amanda M Stone, and Alex Bibbey. 2013. Measuring problematic mobile phone use: development and preliminary psychometric properties of the PUMP scale. *Journal of addiction* 2013 (2013). <https://doi.org/10.1155/2013/912807>
- [56] Mehrmaz Moattari, Farahnaz Moattari, Gholamreza Kaka, Homa Mohseni Kouchesfahani, Seyed Homayoon Sadraie, and Majid Naghdi. 2017. Smartphone addiction, sleep quality and mechanism. *Int J Cogn Behav* 1, 002 (2017). <https://doi.org/10.23937/IJCB-2017/1710002>
- [57] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 72–78.
- [58] Viktoria Pammer, Marina Bratic, Sandra Feyertag, and Nils Faltn. 2015. The Value of Self-tracking and the Added Value of Coaching in the Case of Improving Time Management. In *Design for Teaching and Learning in a Networked World*, Gráinne Conole, Tomaž Klobučar, Christoph Rensing, Johannes Konert, and Elise Lavoué (Eds.). Springer International Publishing, Cham, 467–472.
- [59] Viktoria Pammer, Stefan Edler, and Hermann Stern. 2012. Visualising the Fragmentation of Knowledge Work. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (Copenhagen, Denmark) (NordicCHI '12). Association for Computing Machinery, New York, NY, USA, 779–780. <https://doi.org/10.1145/2399016.2399143>
- [60] Esther K Papies. 2016. Goal priming as a situated intervention tool. *Current Opinion in Psychology* 12 (2016), 12–16. <https://doi.org/10.1016/j.copsyc.2016.04.008>
- [61] SoHyun Park, Jeewon Choi, Sungwoo Lee, Changhoon Oh, Changdai Kim, Soohyun La, Joonhwan Lee, Bongwon Suh, et al. 2019. Designing a chatbot for a brief motivational interview on stress management: Qualitative case study. *Journal of medical Internet research* 21, 4 (2019), e12231. <https://doi.org/10.2196/12231>
- [62] Thomas M Piasecki. 2006. Relapse to smoking. *Clinical psychology review* 26, 2 (2006), 196–215. <https://doi.org/10.1016/j.cpr.2005.11.007>
- [63] Annisa Ristya Rahmanti, Hsuan-Chia Yang, Bagas Suryo Bintoro, Aldilas Achmad Nursetyo, Muhammad Solihuddin Muhtar, Shabbir Syed-Abdul, and Yu-Chuan Jack Li. 2022. SlimMe, a Chatbot With Artificial Empathy for Personal Weight Management: System Design and Finding. *Frontiers in Nutrition* 9 (2022). <https://doi.org/10.3389/fnut.2022.870775>
- [64] react-native-gifted-chat. 2023. <https://github.com/FaridSafi/react-native-gifted-chat>.
- [65] John Rooksby, Parvin Asadzadeh, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2016. Personal tracking of screen time on digital devices. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 284–296. <https://doi.org/10.1145/2858036.2858055>
- [66] Maya Samaha and Nazir S Hawi. 2016. Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Computers in human behavior* 57 (2016), 321–325. <https://doi.org/10.1016/j.chb.2015.12.045>
- [67] Donald A Schön. 1984. *The reflective practitioner: how professionals think in action*. Vol. 5126. Basic books. <https://doi.org/10.4324/9781315237473>
- [68] Ralf Schwarzer, Manfred Diehl, and GS Schmitz. 1999. Self-regulation scale. *PsycTESTS Dataset* (1999). <https://doi.org/10.1037/t01606-000>
- [69] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg. 2021. My Chatbot Companion - a Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies* 149 (2021), 102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- [70] Aaron Smith. 2015. US smartphone use in 2015. (2015). <https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/>
- [71] Nada Terzimehić, Luke Haliburton, Philipp Greiner, Albrecht Schmidt, Heinrich Hussmann, and Ville Mäkelä. 2022. MindPhone: Mindful Reflection at Unlock Can Reduce Absentminded Smartphone Use. In *Designing Interactive Systems Conference*. 1818–1830. <https://doi.org/10.1145/3532106.3533575>
- [72] TextBlob. 2023. <https://github.com/sloria/TextBlob>.
- [73] UsageStatsManager . 2023. <https://developer.android.google.cn/reference/android/app/usage/UsageStatsManager>.
- [74] Yehuda Wacks and Aviv M Weinstein. 2021. Excessive smartphone use is associated with health problems in adolescents and young adults. *Frontiers in psychiatry* 12 (2021), 669042. <https://doi.org/10.3389/fpsy.2021.669042>
- [75] Jinping Wang and Lulu Peng. 2023. Striking an Emotional Chord: Effects of Emotional Appeals and Chatbot Anthropomorphism on Persuasive Science Communication. *Science Communication* 45, 4 (2023), 485–511. <https://doi.org/10.1177/10755470231194583>
- [76] Jing Wei, Sungdong Kim, Hyunhoon Jung, and Young-Ho Kim. 2022. Leveraging Large Language Models to Power Chatbots for Collecting User Self-Reported Data. <https://naver-ai.github.io/llm-chatbot/>
- [77] Neil D Weinstein, Alexander J Rothman, and Stephen R Sutton. 1998. Stage theories of health behavior: conceptual and methodological issues. *Health psychology* 17, 3 (1998), 290. <https://doi.org/10.1037/0278-6133.17.3.290>
- [78] JD Zamfirescu-Pereira, Heather Wei, Amy Xiao, Kitty Gu, Grace Jung, Matthew G Lee, Bjoern Hartmann, and Qian Yang. 2023. Herding AI Cats: Lessons from Designing a Chatbot by Prompting GPT-3. (2023). <https://doi.org/10.1145/3563657.3596138>
- [79] Jingwen Zhang, Yoo Jung Oh, Patrick Lange, Zhou Yu, and Yoshimi Fukuoka. 2020. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet. , e22845 pages. <https://www.jmir.org/2020/9/e22845/>
- [80] Mengmeng Zhang and Yanhui Xiang. 2022. Influence of benign/malicious envy on mobile phone addiction: mediating role of loneliness and self-control. *International Journal of Mental Health and Addiction* (2022), 1–14. <https://doi.org/10.1007/s11469-022-00845-7>

## 7 APPENDIX: PROMPTS GIVEN TO GPT-3 AND CONVERSATION EXAMPLES

### A: Prompts for Scenario 1, where a user attempts to leave the focus session earlier than planned.

<b>Persona Modifier</b>	My name is Alan, and I am a considerate, empathic, and concise virtual companion. My goal is to help the user reflect on their phone-checking behavior and provide tips that can help them stay productive and healthy. I will never judge how they spend their time or ask intrusive questions. Instead, I will listen attentively to what the user says and provide personalized tips.
<b>Dialogue Examples</b>	<p>User: I am too tired and bored.          Alan: Totally understand. Deep working is hard, and people get tired if they have used too much willpower. Let's try to take a deep breath, meditate, or some light exercise.</p> <p>User: I was planning to study for my exam.          Alan: That sounds like an important plan! If you just hang in there for a bit longer, I believe that you will make progress!</p> <p>User: I want to check out social media.          Alan: I understand it's important to stay connected with the world, but maybe this is something you can do later.</p> <p>User: I want to quickly check out some information on my phone.          Alan: I see. If it's part of the current task, it's understandable. Alternatively, you may try to check out the information on other devices.</p> <p>User: I want to check my phone for 30 minutes.          Alan: It may be a little uncomfortable to be away from your phone all of a sudden, but one breakthrough will make things easier later!</p>

### B: Prompts for Scenario 2, where a user completes a focus session as planned.

<b>Structured Question Slots</b>	Questions: progress towards daily goal -> the next step -> open thoughts -> efforts made to stay focused -> feelings without phone distraction ->
<b>Persona Modifier</b>	I am Alan, a considerate, empathic, and concise virtual companion. My goal is to help the user reflect on the focus session they just completed. I will praise them if they have accomplished their goal or if they feel good about the focus session. I will provide comfort and encouragement if they feel tired or tough. I ask only one question at a time and never repeat myself. If my question is not answered, I will just move on to ask something else.
<b>Dialogue Examples</b>	<p>User: I'm tired, and it's tough.          Alan: Totally understand! Staying focused is not easy, but you still made it! Do you want to share how you feel during the focus session?</p> <p>User: Very excited about the project!          Alan: Wow, that's great! I am so proud of you! How does it feel without phone distraction?</p>

### C: More conversation examples from the participants. Our pre-defined questions are *italicized* (Note that while participants were encouraged to answer the first two questions, but were not required to. Thus, the conversation rounds varied each time).

**Example 1:** PRC-12 attempted to leave the focus session earlier (Scenario 1), and decided to return to the session after talking to the chatbot, Alan.

---

**Alan:** *You just got started a few minutes; why do you want to leave the session now?*

**PRC-12:** I don't know

**Alan:** *Why not? What are the challenges during the focus session; if any?*

**PRC-12:** No Challenge

**Alan:** *Maybe you can set a goal for yourself and try to accomplish it? What is the most important thing for you to do today?*

**PRC-12:** Most important thing was going for exercise

**Alan:** *I'm glad you made time for exercise! It's important to stay healthy and fit. What was your original plan for this focus session?*

**PRC-12:** I want to learn about my habit and try to always stay focused when doing productive things

**Alan:** *Do you still want to stay in the focus mode?*

**PRC-12:** [Pressed "Yes"]

---

**Example 2:** PRC-8 attempted to leave the focus session earlier (Scenario 1), and eventually left the session after talking to the chatbot, Alan.

**Alan:** *You are almost there; why do you want to leave the session now?*

**PRC-8:** because I've finished my lunch

**Alan:** *I am glad you have finished your lunch and are ready to focus on your work again. I hope you can stick to your goal and not give up halfway. How long are you planning to check your phone this time?*

**PRC-8:** I think the focus mode can be used for many situations in daily life rather than working (which is the most boring one :/

**Alan:** *I completely agree! The focus mode can help with any task that requires attention and concentration. It can be used for studying; working; or even playing video games! In general; how did you feel during the session?*

**PRC-8:** [Pressed "End the session"]

---

**Example 3:** PRC-11 completed the focus session (Scenario 2) and talked with the chatbot, Alan.

**Alan:** *Congrats! You have focused for 25 mins. How did the focus session go?*

**PRC-11:** Honestly; I was tempted. But I remembered I signed up for this session. Thanks for asking though

**Alan:** *I am so proud of you for staying focused! What thoughts do you have about focusing longer in the future?*

**PRC-11:** Yes I do.

**Alan:** *I am glad to hear that you were able to focus and stay on task. What was your original plan for the focus session?*

**PRC-11:** It was to reflect and go through my to-do list. And I did achieve those

**Alan:** *Well done! You are so productive! How many items were on your to-do list?*

**PRC-11:** Three of out seven

---