Jurnal Personalia Pelajar 29(1): 345-357
https://doi.org/10.17576/personalia.2901.2026.36
ISSN 0128-2735

Notifactalk: Designing A Notification System For Overcoming Procrastination Using A Voice-Based Conversational Agent

(Notifactalk: Mereka Bentuk Sistem Pemberitahuan Untuk Mengatasi Prokrastinasi Menggunakan Ejen Perbualan Berasaskan Suara)

TAICHI INOHARA, ATSUSHI HANDA & IKKAKU KAWAGUCHI

ABSTRACT

Procrastination is increasingly understood as a failure of emotion regulation, where individuals avoid tasks to escape negative feelings. Traditional asynchronous reminders, such as push notifications, lack the "forcefulness" to break this inertia and are easily ignored. Conversely, while human-led interventions like phone calls are effective, they are limited by high social costs and scalability issues.To address this, we propose Notifactalk, a novel intervention method that integrates Synchronous Intrusiveness with Adaptive Scaffolding. The system triggers a simulated incoming call using the smartphone's native telephony interface at a scheduled time, forcing an immediate decision to answer. Once connected, a generative AI agent engages the user in a real-time voice dialogue, providing emotional validation, task decomposition, and encouragement. This approach aims to lower psychological barriers and strongly prompt task initiation. This paper details the system design of Notifactalk and outlines an experimental plan to evaluate its effectiveness in mitigating procrastination compared to text-based and one-way auditory interventions.

Keywords: Procrastination, Human-AI Interaction, Interaction Design, Behavior Change

ABSTRAK

Penangguhan kerja kini semakin difahami sebagai satu kegagalan regulasi emosi, di mana individu mengelakkan tugasan bagi menjauhi perasaan negatif. Peringatan tak segerak tradisional, seperti notifikasi tolak, kekurangan "daya desakan" untuk memecahkan inersia ini dan sering diabaikan. Sebaliknya, walaupun intervensi berasaskan manusia seperti panggilan telefon didapati berkesan, ia mempunyai kekangan dari segi kos sosial yang tinggi dan isu kebolehskalaan. Bagi menangani isu ini, kami mencadangkan NotifAcTalk, sebuah kaedah intervensi baharu yang menggabungkan Gangguan Segerak dengan Perancahan Adaptif. Sistem ini mencetuskan panggilan masuk simulasi menggunakan antara muka telefoni asli telefon pintar pada waktu yang telah dijadualkan, sekaligus memaksa pengguna membuat keputusan segera untuk menjawab. Sebaik sahaja disambungkan, ejen AI generatif akan berinteraksi dengan pengguna melalui dialog suara masa nyata, memberikan validasi emosi, penguraian tugasan, dan kata-kata semangat. Pendekatan ini bertujuan untuk merendahkan halangan psikologi dan mendorong permulaan tugasan secara berkesan. Kertas kerja ini memperincikan reka bentuk sistem NotifAcTalk dan menggariskan pelan eksperimen bagi menilai keberkesanannya dalam mengurangkan penangguhan kerja berbanding intervensi berasaskan teks dan audio sehala.

Kata Kunci: Penangguhan Kerja, Interaksi Manusia-AI, Reka Bentuk Interaksi, Perubahan Tingkah Laku

INTRODUCTION

The act of delaying tasks that must be accomplished is called "procrastination" (Schraw et al., 2007; Steel, 2007). It not only negatively affects mental and physical well-being (Johansson et al., 2023; F. M. Sirois, 2023; F. Sirois & Pychyl, 2013a) but also leads to a decline in both personal and organizational productivity(d'Abate & Eddy, 2007; Metin et al., 2016). Despite its negative consequences, procrastination remains pervasive. Research indicates that 80% to 95% of university students have procrastinated at some point (Steel, 2007). Even when people have a clear intention to perform a task, they often end up procrastinating (Rothman et al., 2011; Sheeran, 2002). Addressing this issue has become increasingly critical due to the rising demand for self-regulation in modern environments, driven by factors such as the spread of short-form videos (Xie et al., 2023), the shift to remote work following the COVID-19 pandemic (Qi et al., 2023), and the rise of ICT in education (Broadbent & Poon, 2015; Hashmi et al., 2025).

Smartphone notifications, such as alarms and calendar reminders, are one of the most common methods for mitigating procrastination. However, they are often insufficient because they rely on one-way communication using only visual displays and monotonous beep sounds, which users tend to easily ignore (Kim & Song, 2024). Consequently, existing notifications are notably lacking in the 'forcefulness' required to prompt users to work on tasks.

Interventions involving human interaction are particularly effective for overcoming procrastination due to their interactive and conversational nature (KIM & FOGG, 1997). Moreover, positive feedback plays a crucial role by enhancing self-efficacy, while the act of publicly declaring goals and sharing progress with others often referred to as public commitment is widely recognized as effective for initiating action and achieving objectives (Bandura, 1977; Hollenbeck et al., 1989; Pereira & Díaz, 2021). In particular, telephone communication is known to foster a stronger sense of social connection compared to text-based methods (Daft & Lengel, 1986; Kumar & Epley, 2021), and its effectiveness has been demonstrated in fields such as medical care (Sugumaran et al., 2024; Unick et al., 2024) and counseling (Shih et al., 2023).

However, asking someone to call just for minor daily activities, such as exercise or household chores, is impractical and socially costly. Furthermore, research suggests that soliciting help from individuals or professionals, such as doctors, can be psychologically stressful and deter individuals from seeking support (Vogel et al., 2006). Therefore, human intervention is not scalable for daily use.

The HCI community has explored various strategies to combat procrastination. Gamification, the application of typical game elements (e.g., point scoring, competition, rules of play) to non-game contexts, is one of the most popular approaches in previous HCI studies for reducing procrastination by motivating users (Hunicke et al., 2004; Kirchner-Krath et al., 2024). However, studies point out that it often leads to counterproductive behaviors because users focus on winning the game rather than the task itself (Baker et al., 2004; Diefenbach & Müssig, 2019; Hanus & Fox, 2015; Toda et al., 2018).

Similarly, approaches leveraging external pressure, such as peer monitoring or financial penalties (Beeminder, n.d.; StickK, n.d.), have proven effective in enforcing immediate compliance. However, these methods often induce high levels of psychological stress and anxiety (R. Ryan & Deci, 2000; R. M. Ryan & Deci, 2000). Such negative emotional responses can undermine motivation, making long-term adherence to the system difficult for users.

As for auditory interventions, Kim and Song demonstrated that utilizing the user's own voice for reminders outperforms standard alarms (Kim & Song, 2024). Moreover, previous research indicates that spoken communication is more effective than text at conveying paralinguistic cues and fostering a sense of social connection (Kumar & Epley, 2021). However, such pre-recorded interventions remain static and one-way. Since pre-recorded messages cannot respond to real-time user feedback, they often fail to provide the necessary scaffolding when users face specific psychological barriers or contextual challenges.

Recent advancements in Large Language Models (LLMs) have enabled chatbots to offer empathetic and context-aware support (Beale et al., 2025; Bhattacharjee et al., 2024). However, existing systems remain predominantly reactive, relying on user-initiated text or voice interactions. This lack of proactivity is particularly problematic for disrupting procrastination, as users must voluntarily initiate the interaction.

Consequently, the passive nature of current interfaces limits their effectiveness as proactive interventions. Therefore, designing an effective intervention for procrastination necessitates a dual approach. First, it requires synchronous and high-intrusiveness interaction. Unlike passive text or monotonous auditory alerts, the intervention must leverage a modality that demands immediate attention and evokes a sense of social presence to effectively break the user's inertia. Second, it necessitates adaptive scaffolding. Since procrastination often stems from emotional regulation failure rather than mere forgetfulness, the system must go beyond static reminders. It needs to provide real-time, context-aware persuasion that lowers the user's psychological hurdles. While existing solutions have addressed these factors in isolation, they have failed to integrate them into a cohesive system that can be seamlessly integrated into users' daily lives.

1. Set Task 2. Receive Call 3. Interact with AI

FIGURE 1. User Interaction Flow with UI Screenshots

To implement this dual approach, we propose Notifactalk, a novel notification method that leverages generative AI to provide interactive voice interventions via simulated phone calls. Specifically, the system utilizes the smartphone's native telephony interface (e.g., iOS CallKit) to generate an incoming call. This design ensures synchronous intrusiveness, making the intervention significantly harder to ignore compared to standard push notifications. Furthermore, Notifactalk incorporates real-time voice dialogue powered by a Large Language Model (LLM). Unlike static scripts, the AI agent dynamically adapts its persuasion strategy - offering empathy, decomposing tasks, and providing encouragement while encouraging autonomous initiation by allowing users to terminate the call once they are ready thereby functioning as a catalyst for behavioral change.

In this paper, we detail the design of Notifactalk based on previous research and describe its implementation. We then propose an experimental design for evaluating its effectiveness. We also discuss the expected implications and insights gained through the system's design process, providing a foundation for future empirical validation.

LITERATURE REVIEW

To situate our work within the broader context of Human-Computer Interaction (HCI) and procrastination research, this section reviews the theoretical mechanisms of procrastination and analyzes why existing technological interventions often fail to induce behavioral change. We then discuss the limitations of "asynchronous" nudges, such as push notifications and gamification. Finally, we examine the potential of synchronous auditory interventions and the role of social presence in interactive systems to establish the theoretical foundation for our proposed system, NotifAcTalk, leading to our research hypotheses and research questions.

Mechanisms of Procrastination and The Limits of Asynchronous Nudges

Procrastination as Emotion Regulation Failure

While procrastination is commonly misidentified as a deficiency in time management or a lack of willpower, recent psychological research characterizes it as a failure of emotion regulation (F. Sirois & Pychyl, 2013b; Solomon & Rothblum, 1984). When individuals are confronted with tasks perceived as difficult, boring, or ambiguous, they experience negative emotions such as anxiety, insecurity, or boredom.

To alleviate this negative mood in the short term, individuals voluntarily delay the task in favor of immediate mood repair, despite knowing that this delay will result in negative consequences later (F. Sirois & Pychyl, 2013b). This mechanism explains why individuals often fail to act even when they possess strong behavioral intentions and clear plans.

The core issue is not that the user forgets to perform the task, but rather that they are actively avoiding the psychological distress associated with it. Consequently, interventions that simply remind users of tasks without addressing these emotional barriers are often insufficient.

The "Asynchronous" Barrier in Existing Interventions

To combat procrastination, the HCI community has developed various technological interventions, ranging from productivity apps (Lu et al., 2014) and to-do lists (Torino et al., 2025) to systems integrating gamification elements (Hunicke et al., 2004; Kirchner-Krath et al., 2024).

Gamification, which applies game design elements (e.g., points, leaderboards) to non-game contexts, aims to boost user motivation. While these approaches have shown promise in enhancing task engagement, users may focus on "gaming the system" rather than the task itself, leading to counterproductive behaviors such as avoiding challenging tasks or falsifying task completion to avoid penalties(Baker et al., 2004; Diefenbach & Müssig, 2019; Hanus & Fox, 2015; Toda et al., 2018). More recently, Large Language Models (LLMs) have been employed to provide personalized task management or redirect users from distractions (Beale et al., 2025; Bhattacharjee et al., 2024).

However, a critical limitation shared by these text-based chatbots, gamified apps, and standard push notifications is that they operate primarily as asynchronous interventions. In an asynchronous interaction, the decision of "when to attend to the message" resides largely with the user. When a notification appears on a screen, it competes with other stimuli and requires the user's voluntary attention and cognitive effort to attend.

For a procrastinator already in a state of avoidance, these cues lack high physical or psychological "intrusiveness," thereby failing to exert the necessary force to disrupt the user's avoidance behavior effectively. To bridge the gap between intention and action, an intervention must therefore possess a synchronous quality that demands immediate attention and cannot be easily disregarded.

Synchronous Intervention

The Necessity of High Intrusiveness

As discussed in the previous section, asynchronous cues are insufficient to overcome the emotional barrier of procrastination. To disrupt the user's avoidance behavior effectively, an intervention must be synchronous and possess a level of intrusiveness that cannot be easily ignored.

In this context, the smartphone's native telephony interface offers a distinct advantage over standard notifications. Technically, an incoming call occupies the entire screen and continues to play a ringing sound until the user takes action, thereby demanding immediate visual and auditory attention.

Sociological studies also suggest that telephone ringtones carry a strong signal of "urgency" and "presence," creating a psychological pressure to respond that is significantly higher than that of text-based alerts (Bednall et al., 2011; Licoppe, 2004; SCHEGLOFF, 1968). Empirical evidence from healthcare domains supports the efficacy of this modality; interventions utilizing synchronous telephone calls have been shown to significantly improve medication adherence (e.g., for diabetes management), facilitate habit formation (e.g., exercise for obesity reduction), and enhance counseling outcomes (e.g., for depression treatment) (Shih et al., 2023).

However, employing humans for minor daily tasks is impractical, socially costly, and not scalable for daily use.

Confounding Factors in Previous Studies

While the efficacy of synchronous interventions such as phone calls is well-documented, a critical gap remains in understanding the specific drivers of this behavioral change. Given that prior research has primarily focused on medical or counseling contexts, existing literature typically evaluates such interventions as a composite whole, conflating the effect of "synchronous interruption" (the act of receiving a call) with the effect of "dialogue content" (the conversation that follows) (Leach & Christensen, 2006; Michie et al., 2011).

Consequently, it remains unclear whether the initial disruption of procrastination is triggered by the sheer force of the interruption which breaks the user's inertia or by the persuasive content provided during the interaction. This study aims to disentangle these two factors. By comparing a non-interactive call intervention utilizing static recorded speech (One-way) against an interactive intervention call (Notifactalk), we seek to isolate the specific contributions of synchronous intrusiveness versus interactive persuasion in the context of procrastination.

Beneficial Discomfort as a Trigger

Introducing such a highly intrusive intervention may raise concerns about user experience. However, contrary to traditional usability paradigms that prioritize frictionless interaction, HCI researchers have increasingly recognized the strategic value of "Uncomfortable User Experience." Benford et al. argue that managed discomfort whether visceral, cultural, or control-related can be employed to enhance engagement, establish stronger user commitment, and facilitate enlightenment (Benford et al., 2013).

This perspective aligns with psychological frameworks suggesting that while positive affect signals a satisfactory status quo, negative emotions (e.g., anxiety or urgency) are more effective at activating problem-solving mindsets (Pekrun, 1992). Similarly, Rekimoto and Tsujita proposed "Inconvenient Interactions," positing that introducing intentional friction can compel users to perform specific actions to resolve the discomfort (Rekimoto & Tsujita, 2014).

Previous studies have demonstrated the utility of this mechanism. For instance, Oh et al. designed an alarm system that requires users to solve math problems to terminate the noise (Oh et al., 2022). Although this induces immediate unpleasantness, it effectively achieves the higher-order goal of waking up.

We posit that the "forced decision point" created by a synchronous phone call operates on this same principle. By generating a high-priority interruption that cannot be passively ignored, the system imposes a "beneficial discomfort" that jolts the user out of the procrastination loop.

From Monologue to Dialogue

The Limits of One-Way Communication

Even if synchronous interventions effectively disrupt avoidance behavior and force a user reaction, securing this initial attention alone does not guarantee task execution. Existing interventions have predominantly relied on one-way communication without any capacity for user feedback. This "monologue" approach often leads users to reflexively dismiss the prompt. Consequently, the "intention-action gap" remains unbridged.

To address this, we propose incorporating a Generative AI-driven interactive approach. By shifting from monologue to dialogue, we aim to transform the system from a passive reminder into an active agent that can adapt its persuasion strategy in real-time.

The Neurocognitive Advantage of Voice

To support this interactive approach, we employ the human voice as the primary medium, leveraging its neurocognitive superiority for capturing attention. Research in cognitive engineering and evolutionary psychology posits that the human auditory system is "tuned" to prioritize vocal stimuli over other environmental sounds.

As demonstrated by Kim and Song, the human voice conveys paralinguistic cues that elicit an instinctive orienting response more effectively than standard auditory alarms (Kim & Song, 2024). This characteristic ensures that the AI's intervention is not merely heard as background noise but is cognitively processed as a signal demanding social attention. Ultimately, while voice ensures the message is received, it is the interaction that ensures the message is acted upon.

Scaffolding and Social Presence

The integration of interactive AI introduces two critical mechanisms for behavioral change: Scaffolding and Social Presence.

First, unlike static notifications, an AI agent can provide Scaffolding - a technique where support is dynamically tailored to the user's capacity (Haykin & Chen, 2005; Nass & Gong, 2000; Ulfvengren, 2003). Since procrastination is typically driven by emotion regulation failure, the AI can directly address this by validating the user's feelings and facilitating task decomposition, such as proposing "just 5 minutes" when the user feels overwhelmed. This allows the system to mitigate the psychological barriers that prevent task initiation, effectively guiding the user through the difficulty.

Second, the interactive nature of the system leverages the Computers As Social Actors (CASA) paradigm (Nass et al., 1994; Reeves & Nass, 1996). Research suggests that humans unconsciously apply social rules to conversational partners, even when they know the partner is a machine, creating a sense of social accountability. This psychological commitment makes it significantly harder to ignore the agent compared to a one-way recording (Kalbach, 2002; Nass & Moon, 2000). Therefore, this study investigates not only the behavioral outcomes but also the psychological impacts of AI-mediated dialogue, hypothesizing that it will alleviate the negative emotions associated with procrastination.

Research Questions and Hypotheses

Based on the theoretical framework established above, this study aims to verify the effectiveness of synchronous interaction and AI-mediated scaffolding in overcoming procrastination. To disentangle the effects of "synchronous intrusiveness" (Section 2.2) and "interactive persuasion" (Section 2.3), we designed a comparative study with three experimental conditions:

i. Baseline (Text): A standard asynchronous push notification.
ii. One-way Voice: A synchronous call that plays a one-way recorded message.
iii. NotifAcTalk (Proposed): A synchronous call with an interactive generative AI agent.

We address the following three Research Questions and corresponding Hypotheses.

RQ1: To what extent does the proposed approach mitigate procrastination behavior?

H1 (Response Time): Due to the high intrusiveness of the telephony channel, both synchronous conditions (One-way Voice and NotifAcTalk) will result in significantly shorter response times compared to the Baseline (Text), thereby verifying the alertness effect of the modality.

H2 (Task Completion): Due to the adaptive persuasion provided by the dialogue, the Notifactalk condition will yield a higher task completion rate and better performance in task execution compared to the One-way Voice and Baseline conditions, verifying the scaffolding effect of the interaction.

RQ2: How does the proposed approach affect the user's psychological state and experience?

H3 (Emotion Regulation): By providing psychological scaffolding and validation, the Notifactalk condition will reduce negative emotions associated with the task compared to the other conditions.

H4 (Social Presence): Based on the CASA paradigm, the Notifactalk condition will elicit a higher sense of Social Presence compared to the One-way Voice condition and the Baseline.

H5 (Intrusiveness): Reflecting the forced decision-point nature of the intervention, both synchronous conditions (One-way Voice and Notifactalk) will be perceived as significantly more intrusive than the Baseline, confirming that the system is indeed functioning as a discomfort trigger.

RQ3: How do users evaluate the system?

H6 (Beneficial Discomfort): Despite the high perceived intrusiveness (H5), the Notifactalk condition will achieve higher system acceptance and perceived utility than the One-way Voice condition. We postulate that users will tolerate the initial discomfort of the call because the subsequent AI support provides sufficient value to justify the interruption.

SYSTEM IMPLEMENTATION

To verify the effectiveness of synchronous intrusiveness and adaptive scaffolding, we designed and implemented Notifactalk, an integrated mobile application. The system initiates interventions via simulated phone calls at scheduled times and guides users through learning tasks using real-time voice dialogue with a generative AI agent. This section details the system architecture and technical implementation constructed to address the design requirements established in Section 2.

System Architecture

Notifactalk is primarily composed of two modules: the Intervention Control Module and the Real-time Dialogue Module. The system achieves personalized interventions by coordinating external server-side trigger management with client-side dialogue processing.

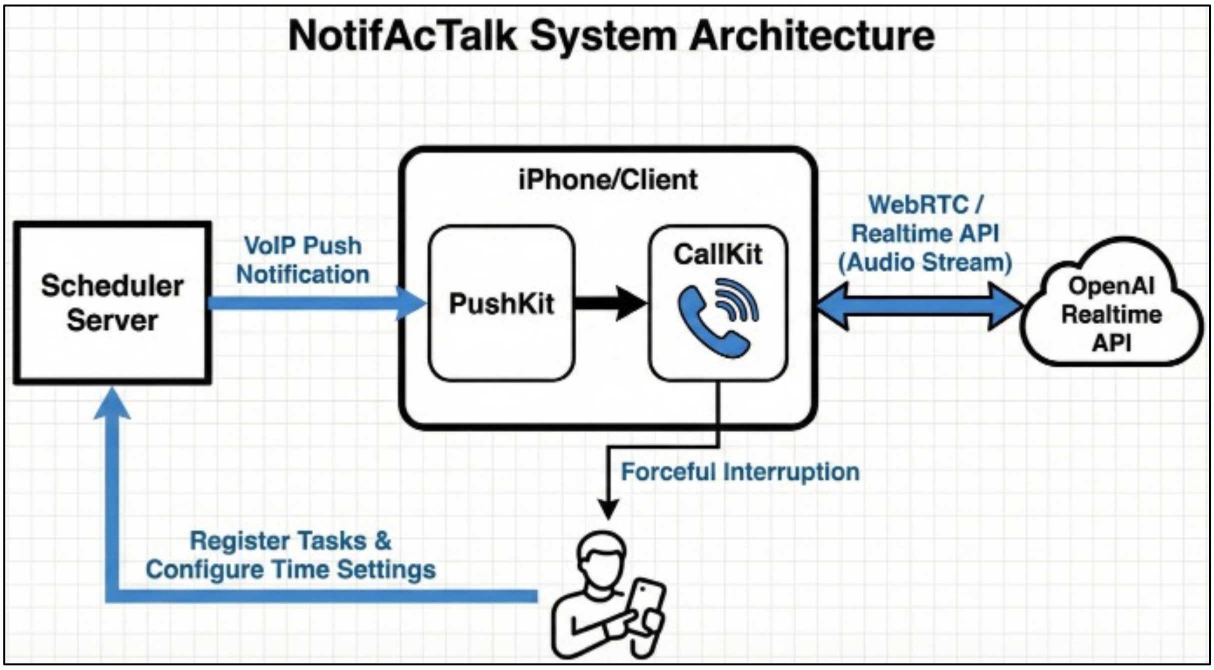

FIGURE 2. NotifAcTalk System Architecture

Flow and Connectivity

The intervention flow begins with the user's input via the application's scheduling interface, allowing for flexible adjustments to the intervention timing. Based on this schedule, an external server transmits a VoIP notification (utilizing PushKit on iOS) to the mobile client. This mechanism ensures that the application can be reliably activated in the background, even when the application is terminated or the device is locked.

Upon receiving the notification, the client leverages the OS-native telephony framework (e.g., CallKit on iOS) to present a full-screen incoming call interface and trigger the ringtone. When the user answers this call, the client immediately establishes a WebRTC connection to the generative AI endpoint (the OpenAI Realtime API). Upon establishing this connection, the system initializes the AI with specific contextual parameters such as persona definition (e.g., an empathetic coach) and the user’s local time. This allows the AI to initiate the dialogue with an understanding of the user's context.

Minimal Latency Architecture

A critical challenge in voice dialogue systems is latency. Conventional architectures typically rely on disjointed pipelines- Automatic Speech Recognition (ASR), Large Language Models (LLM), and Text-to-Speech (TTS) which introduce delays that disrupt natural turn-taking.

To maximize Social Presence, we adopted a streamlined architecture that establishes a direct WebRTC connection from the mobile client to the AI model. By processing audio input and output within a single multimodal model (via the OpenAI Realtime API), this approach eliminates the latency overhead associated with discrete processing stages. Consequently, the system achieves low-latency response times comparable to natural human conversation, facilitating smooth and spontaneous interaction.

Telephony Interface

This section describes the client-side implementation designed to establish "Synchronous Intrusiveness" (as defined in Section 2) and seamlessly transition the user into an external task environment.

Native Call Emulation

A defining feature of this system, distinguishing it from standard text notifications, is the utilization of the OS-native telephony UI for interruption. By leveraging telephony frameworks (CallKit on iOS), the system emulates the exact behavior of a real incoming phone call. This ensures that a full-screen call interface appears and the ringtone persists, even if the device is locked or the user is engaged in other applications. This design imposes a "Forced Decision Point," compelling the user to make an immediate binary choice - answer or decline thereby breaking the inertia of procrastination (addressing H1 and H5).

Persistent Audio for Scaffolding

To verify the effectiveness of scaffolding (task breakdown and guidance), this study employs an external web-based learning assignment as the target task. NotifAcTalk supports persistent background VoIP connectivity, enabling the system to accompany the user across application boundaries.

Once the user answers the call, the audio connection is maintained even as the user minimizes the app to open a web browser and access the external learning site which involves multiple steps such as logging in. This "Persistent Audio" mechanism allows the AI to provide verbal guidance - such as negotiating the first step ("Let's just open the browser first") or checking status ("Have you logged in?") without occupying the visual interface. This is critical for preventing attrition during high-friction transition stages, such as switching applications or setting up the study environment.

Real-time Dialogue and Adaptive Scaffolding

A critical distinction between this system and traditional one-way monologues is the user's ability to interrupt at any moment. Even mid-sentence, if the user interjects for example, by saying "Wait a minute" the AI immediately halts its output to listen, mimicking natural human turn-taking dynamics by leveraging the Voice Activity Detection (VAD) capabilities of the OpenAI Realtime API. This capability fosters a heightened sense of Social Presence (H4), serving as the foundation for engaging the user in a genuine dialogue.

Within this interactive framework, the AI is prompted to function not merely as a chatbot, but as an empathetic partner. The system analyzes the context of the user's utterances and emotional state to dynamically select the most appropriate persuasion strategy. Primary strategies include Emotional Validation, where the AI acknowledges rather than dismisses negative user reactions (e.g., validating "I am tired" with empathy), thereby lowering psychological barriers. This is complemented by Task Decomposition; when the user is unable to initiate action, the AI negotiates manageable micro-tasks such as proposing, "Let's just open the browser for now" to facilitate agreement on task initiation.

Crucially, the termination of the dialogue is designed to be user-initiated. While the AI's primary goal is to guide the user to start the task, the connection is maintained until the user feels sufficiently ready to proceed independently. By granting the user the authority to end the call, the system encourages a psychological transition from "being supported" to "autonomous execution." This design ensures that the scaffolding is removed only when the user has internally crossed the threshold of initiation.

All dialogue logs, including the duration of the call until user-disconnection, and subsequent task latencies are automatically recorded to serve as the basis for the analysis detailed in Section 4.

EXPERIMENTAL DESIGN

This section outlines the comparative experiment planned to evaluate the effectiveness of NotifAcTalk. The primary objective is to empirically validate the Research Questions (RQ1–RQ3) and hypotheses (H1–H6) established in Section 2. The experiment is scheduled to commence in early January 2026.

Overview

To isolate and evaluate the effects of "Intrusiveness" and "Interactivity" on behavioral change and psychological states, we employ a between-subjects design with three conditions. Participants will be randomly assigned to one of the following three conditions:

Text Notification Group (Baseline)

This group serves as the control condition, where participants receive standard text-based push notifications. This condition represents a (Low Intrusiveness, No Interactivity) environment, establishing the baseline for evaluating existing reminder methods and providing a reference for measuring the intrusiveness and acceptability of the voice-based interventions.

One-way Voice Group

In this condition, the intervention occurs via a phone call utilizing CallKit, representing a (High Intrusiveness, No Interactivity) environment. Upon answering the call, a one-way recorded message (e.g., "It is time for your study session. Please open the web page.") is played before the call terminates. This group is designed to isolate the effect of interactivity; by comparing this group with Notifactalk, we intend to verify whether behavioral changes and psychological burden reduction are driven by interactive scaffolding or merely by the alerting effect of the ringtone.

NotifAcTalk Group (Proposed)

This group represents the (High Intrusiveness, High Interactivity) environment. Upon answering the call, participants engage in a real-time dialogue with the generative AI. As detailed in Section 3.3, the AI provides adaptive scaffolding through emotional validation and task decomposition. The connection is maintained until the user decides they are ready to proceed autonomously and manually terminates the call. This condition is intended to demonstrate that the synergy between synchronous intrusiveness and interactive support achieves superior behavioral outcomes while maintaining user acceptance.

Participants

We plan to recruit approximately 18 university and graduate students (6 per condition). To meet the technical requirements for CallKit and PushKit, participants must be regular iPhone users. While a self-identified "habitual procrastinator" status is not a mandatory requirement, we recognize that individual traits significantly influence outcomes. Therefore, we will measure each participant's baseline procrastination tendency using the Pure Procrastination Scale (PPS) in a pre-study survey. In the subsequent analysis, these PPS scores will be utilized as a covariate to adjust for individual differences and ensure a rigorous evaluation of the system's effectiveness.

Task and Procedure

This section details the specific content of the experimental task and the sequential procedure from intervention to data collection.

Learning Task and Performance Measurement

The task for this experiment is to complete a learning session using a custom web-based English vocabulary learning system. The content is divided into multiple sections, each consisting of approximately 15 words; participants are encouraged to complete one section per day. Each section is structured into two stages: a Learning Phase for reviewing word meanings and an Evaluation Phase to measure comprehension. To quantitatively evaluate how the improvement in task initiation contributes to long-term vocabulary retention, we will conduct a diagnostic proficiency test on the first day (Pre-test) and the final day (Post-test) of the experiment. This allows us to verify the system's effectiveness not only in triggering behavior but also in enhancing educational outcomes.

Experimental Procedure

The experiment is scheduled to last for 14 days and will follow the procedure outlined below:

i. Preparation (Pre-study): Participants complete a pre-study survey (e.g., measuring the PPS) and install the dedicated application. Simultaneously, they take the Pre-test to establish a baseline. Intervention times can be flexibly configured by the users within the app to align with their daily schedules.
ii. Daily Intervention: At the scheduled time, the intervention corresponding to the assigned group occurs. The Text Notification group receives a push notification, while the One-way Voice group receives a phone call with a recorded message. In the NotifAcTalk group, a real-time dialogue with the AI takes place to provide motivation and guidance, aimed at lowering psychological barriers to task initiation.
iii. Task Execution and Post-task Survey: Participants perform the daily learning and testing via a web browser. Immediately following the quiz, an in-situ survey employing the Experience Sampling Method (ESM) is presented. This captures the participant's emotional state and impressions of the system in real-time. This integrated design aims to maximize response rates while capturing psychological data immediately after the learning activity.
iv. Post-study: After the 14-day period, participants take the Post-test, complete a final survey regarding system acceptance (e.g., the Technology Acceptance Model, TAM), and participate in a semi-structured interview to investigate qualitative changes in their interaction experience.

Measures

This study employs a mixed-methods approach, combining both quantitative and qualitative data to evaluate the impact of the proposed system on mitigating procrastination, reducing psychological burden, and ensuring acceptability. The evaluation metrics are categorized into four dimensions: behavior, learning, psychology, and acceptance.

Quantitative Metrics

In the quantitative evaluation, we will first record behavioral data to measure the direct effect of the intervention: Response Latency [H1], defined as the duration from the initiation of the intervention to task commencement (opening the web page), and the Completion Rate [H2] over the 14-day experimental period. Additionally, to verify whether the intervention encourages genuine engagement rather than mere compliance, we will measure Learning Performance (accuracy on daily quizzes). Furthermore, by comparing scores from the Pre-test and the Post-test, we will calculate the Learning Gain to analyze how the nature of the intervention influences knowledge retention.

Changes in participants' subjective psychological states will be captured through a combination of daily in-situ surveys and a post-study questionnaire. To prevent attrition due to response fatigue, the daily Experience Sampling Method (ESM) surveys conducted immediately after each task will be limited to three key items: immediate emotional state (measuring negative affect via a shortened PANAS scale) [H3], the momentary sense of Social Presence [H4], and the Perceived Intrusiveness of the intervention [H5]. As a longitudinal evaluation, the post-study questionnaire will collect data on overall Social Presence and anthropomorphism to measure the extent to which participants perceived the AI as a social entity, along with the Technology Acceptance Model (TAM) [H6] to assess perceived usefulness and intention to continue using the system.

Qualitative Evaluation

To elucidate the subjective transformation of the user experience, we will conduct semi-structured interviews following the experiment. These interviews will investigate how the highly intrusive phone-call intervention is perceived by users following the interactive process with the AI. Specifically, we will explore whether the AI's emotional validation and verbal cues contributed to lowering psychological barriers and identify the factors behind the transition toward autonomous task initiation (i.e., the user's choice to terminate the call). Thematic Analysis will be applied to the gathered verbal data to extract the underlying psychological mechanisms supporting behavioral change. Furthermore, by utilizing the Pure Procrastination Scale (PPS) as a covariate, we plan to examine how individual procrastination traits influence the perception of the intervention and the quality of learning behavior.

CONCLUSION AND FUTURE WORK

Conclusion

In this study, we proposed Notifactalk, a novel intervention method designed to facilitate task initiation by addressing procrastination as a failure of emotion regulation. The core of this research lies in the conceptualization and implementation of two distinct yet synergistic factors: Synchronous Intrusiveness and Adaptive Scaffolding. The first factor, Intrusiveness, leverages the smartphone's native telephony interface to behaviorally and psychologically disrupt a user's avoidance behavior. This serves as a potent trigger that overcomes the inherent limitations of standard notifications, which are frequently ignored or habituated. The second factor, Scaffolding, utilizes real-time voice dialogue with a generative AI to mitigate psychological barriers and reduce the immediate cognitive load of the task.

To evaluate the individual contributions of these factors, this paper presents an experimental design comparing NotifAcTalk against standard text notifications and a "One-way Voice" condition, which isolates the effect of intrusiveness from interactive support. Through the experiment scheduled for January 2026, we intend to demonstrate the impact of these elements on mitigating procrastination and reducing psychological distress, thereby establishing a new intervention model for self-regulation challenges in the digital age.

Future Work

The framework of Notifactalk proposed in this study holds significant potential for further development. In the future, there is substantial room to refine the quality of the voice and dialogue that characterize the system. For instance, it would be a compelling challenge to explore how user acceptance and social presence change when the agent's persona and voice are personalized allowing users to select voices ranging from their own or a close friend's to an authoritative tone, such as that of a supervisor. Furthermore, by enabling the AI to maintain a long-term memory of previous interactions, the system could provide more personal and deeply grounded support, referencing the user's specific past successes or struggles during the persuasion process.

Another important step is a comparative validation between AI-mediated interventions and actual human-led interventions. While humans exert strong social pressure, AI may offer a unique advantage by providing a "safe space" for consultation, free from the fear of social evaluation or being judged by others. Clarifying this trade-off will allow us to define the unique role of AI agents in behavioral support.

Moreover, the system could be expanded into more proactive and multimodal dimensions. Potential directions include dynamic intervention methods that detect signs of procrastination such as aimless app usage based on smartphone activity logs, enabling the system to determine the optimal timing for intervention based on the user's state. Additionally, integrating visual scaffolding, such as automatically displaying the first step of a task on the screen in synchronization with voice guidance, could further lower the friction of task initiation.

Ultimately, the goal of this research extends beyond merely reducing procrastination. Through technological interventions that nudge users toward the "first step" of an action, we aim to construct a social environment where individuals can engage with their goals autonomously and positively. We envision that the accumulation of small achievements gained through these interventions will enhance users' self-efficacy, ultimately contributing to a society where individuals can overcome temptation and achieve their goals with agency.

REFERENCES

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the cognitive tutor classroom: When students “game the system.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 383–390. https://doi.org/10.1145/985692.985741

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191.

Beale, S. K., Cohen, N., Secheli, B., McIntire, D., & Kho, K. A. (2025). Comparing physician and artificial intelligence chatbot responses to posthysterectomy questions posted to a public social media forum. *AJOG Global Reports*, *5*(3), 100553. https://doi.org/10.1016/j.xagr.2025.100553

Bednall, D., Totten, J., & Lipscomb, T. (2011). *Mobile Phone Etiquette: A Cross-national Comparison*. Beeminder. (n.d.). *Beeminder*. https://www.beeminder.com/

Benford, S., Greenhalgh, C., Giannachi, G., Walker, B., Marshall, J., & Rodden, T. (2013). Uncomfortable user experience. *Commun. ACM*, *56*(9), 66–73. https://doi.org/10.1145/2500468.2500889

Bhattacharjee, A., Zeng, Y., Xu, S. Y., Kulzhabayeva, D., Ma, M., Kornfield, R., Ahmed, S. I., Mariakakis, A., Czerwinski, M. P., Kuzminykh, A., Liut, M., & Williams, J. J. (2024). Understanding the Role of Large Language Models in Personalizing and Scaffolding Strategies to Combat Academic Procrastination. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3613904.3642081

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13.

d’Abate, C. P., & Eddy, E. R. (2007). Engaging in personal business on the job: Extending the presenteeism construct. *Human Resource Development Quarterly*, *18*(3), 361–383.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, *32*(5), 554–571.

Diefenbach, S., & Müssig, A. (2019). Counterproductive effects of gamification: An analysis on the example of the gamified task manager Habitica. *International Journal of Human-Computer Studies*, *127*, 190–210.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152–161.

Hashmi, Z. F., Iqbal, J., Asghar, M. Z., & Siming, L. (2025). The influence of online learning interactions on self-regulated learning: Mediating role of technology proficiencies among higher education students. *Open Learning: The Journal of Open, Distance and e-Learning*, 1–26.

Haykin, S., & Chen, Z. (2005). The Cocktail Party Problem. *Neural Computation*, *17*, 1875–1902. https://doi.org/10.1162/0899766054322964

Hollenbeck, J. R., Williams, C. R., & Klein, H. J. (1989). An empirical examination of the antecedents of commitment to difficult goals. *Journal of Applied Psychology*, *74*(1), 18.

Hunicke, R., LeBlanc, M., Zubek, R., & others. (2004). MDA: A formal approach to game design and game research. *Proceedings of the AAAI Workshop on Challenges in Game AI*, *4*(1), 1722.

Johansson, F., Rozental, A., Edlund, K., Côté, P., Sundberg, T., Onell, C., Rudman, A., & Skillgate, E. (2023). Associations between procrastination and subsequent health outcomes among university students in Sweden. *JAMA Network Open*, *6*(1), e2249346–e2249346.

Kalbach, J. (2002). Persuasive technology: Using computers to change what we think and do. *Ubiquity*, *2002*, 5. KIM, E.-Y., & FOGG, B. (1997). Computers Are Social Actors: A Review of Current. *Human Values and the Design of Computer Technology*, *72*, 137.

Kim, J., & Song, H. (2024). My voice as a daily reminder: Self-voice alarm for daily goal achievement. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–16.

Kirchner-Krath, J., Schmidt-Kraepelin, M., Schöbel, S., Ullrich, M., Sunyaev, A., & Von Korflesch, H. F. (2024). Outplay your weaker self: A mixed-methods study on gamification to overcome procrastination in academia. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–19.

Kumar, A., & Epley, N. (2021). It's surprisingly nice to hear you: Misunderstanding the impact of communication media can lead to suboptimal choices of how to connect with others. *Journal of Experimental Psychology: General*, *150*(3), 595.

Leach, L., & Christensen, H. (2006). A systematic review of telephone-based interventions for mental disorders. *Journal of Telemedicine and Telecare*, *12*, 122–129. https://doi.org/10.1258/135763306776738558 Licoppe, C. (2004). 'Connected' Presence: The Emergence of a New Repertoire for Managing Social Relationships in a Changing Communication Technoscape. *Environment and Planning D-Society & Space - ENVIRON PLAN D-SOC SPACE*, *22*, 135–156. https://doi.org/10.1068/d323t

Lu, Z., Matthews, B., & Viller, S. (2014). Designing for reducing procrastination on side projects. *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, 396–399. https://doi.org/10.1145/2686612.2686673

Metin, U. B., Taris, T. W., & Peeters, M. C. (2016). Measuring procrastination at work and its associated workplace aspects. *Personality and Individual Differences*, *101*, 254–263.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1). https://doi.org/10.1186/1748-5908-6-42

Nass, C., & Gong, L. (2000). Speech interfaces from an evolutionary perspective. *Commun. ACM*, *43*, 36–43. Nass, C., & Moon, Y. (2000). *Machines and Mindlessness*. https://api.semanticscholar.org/CorpusID:15851410 Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. https://api.semanticscholar.org/CorpusID:2739302

Oh, K. T., Ko, J., Shin, J., & Ko, M. (2022). Using Wake-Up Tasks for Morning Behavior Change: Development and Usability Study. *JMIR Formative Research*, *6*. https://api.semanticscholar.org/CorpusID:251946556

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, *41*, 359–376.

Pereira, J., & Díaz, Ó. (2021). Struggling to keep tabs on capstone projects: A chatbot to tackle student procrastination. *ACM Transactions on Computing Education (TOCE)*, *22*(1), 1–22.

Qi, L., Xu, Y., & Liu, B. (2023). Work out of office: How and when does employees' self-control influence their remote work effectiveness? *Frontiers in Psychology*, *14*, 1265593.

Reeves, B., & Nass, C. (1996). *The media equation—How people treat computers, television, and new media like real people and places*. https://api.semanticscholar.org/CorpusID:265223

Rekimoto, J., & Tsujita, H. (2014). Inconvenient interactions: An alternative interaction design approach to enrich our daily activities. *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. https://api.semanticscholar.org/CorpusID:33209595

Rothman, A. J., Baldwin, A. S., Hertel, A. W., & Fuglestad, P. T. (2011). *Self-regulation and behavior change: Disentangling behavioral initiation and behavioral maintenance.*

Ryan, R., & Deci, E. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*, 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020 SCHEGLOFF, E. A. (1968). Sequencing in Conversational Openings1. *American Anthropologist*, *70*(6), 1075–1095. https://doi.org/10.1525/aa.1968.70.6.02a00030

Schraw, G., Wadkins, T., & Olafson, L. (2007). Doing the things we do: A grounded theory of academic procrastination. *Journal of Educational Psychology*, *99*(1), 12.

Sheeran, P. (2002). Intention—Behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, *12*(1), 1–36.

Shih, Y.-H., Wang, J.-Y., Chou, P.-H., & Lin, K.-H. (2023). The effects of treatment via telemedicine interventions for patients with depression on depressive symptoms and quality of life: A systematic review and meta-analysis. *Annals of Medicine*, *55*(1), 1092–1101.

Sirois, F. M. (2023). Procrastination and stress: A conceptual review of why context matters. *International Journal of Environmental Research and Public Health*, *20*(6), 5031.

Sirois, F., & Pychyl, T. (2013a). Procrastination and the priority of short-term mood regulation: Consequences for future self. *Social and Personality Psychology Compass*, *7*(2), 115–127.

Sirois, F., & Pychyl, T. (2013b). Procrastination and the Priority of Short-Term Mood Regulation: Consequences for Future Self. *Social and Personality Psychology Compass*, *7*, 115–127. https://doi.org/10.1111/spc3.12011

Solomon, L. J., & Rothblum, E. D. (1984). Academic procrastination: Frequency and cognitive-behavioral correlates. *Journal of Counseling Psychology*, *31*(4), 503–509. https://doi.org/10.1037/0022-0167.31.4.503

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*, *133*(1), 65.

StickK. (n.d.). *StickK*. https://www.stickk.com/

Sugumaran, T., LeGris, J., Strachan, P. H., Alliston, P., & Sherifali, D. (2024). Experience of telephone-based diabetes health coaching among community-based adults with type 2 diabetes. *Diabetes Spectrum*, *37*(4), 360–368.

Toda, A., Valle, P. H., & Isotani, S. (2018). *The Dark Side of Gamification: An Overview of Negative Effects of Gamification in Education*. https://doi.org/10.1007/978-3-319-97934-2_9

Torino, M., Pastrana, C., & Gatdula, M. (2025). MañanaList App: A Mobile To-Do List Application for Efficient and Personalized Task Management. *International Journal of Research and Innovation in Social Science*, *IX*, 2507–2523. https://doi.org/10.47772/IJRISS.2025.909000215

Ulfvengren, P. (2003). *Design of Natural Warning Sounds in Human-Machine Systems*. https://api.semanticscholar.org/CorpusID:110738392

Unick, J. L., Pellegrini, C. A., Dunsiger, S. I., Demos, K. E., Thomas, J. G., Bond, D. S., Lee, R. H., Webster, J., & Wing, R. R. (2024). An adaptive telephone coaching intervention for patients in an online weight loss program: A randomized clinical trial. *JAMA Network Open*, *7*(6), e2414587–e2414587.

Vogel, D. L., Wade, N. G., & Haake, S. (2006). Measuring the self-stigma associated with seeking psychological help. *Journal of Counseling Psychology*, *53*(3), 325.

Xie, J., Xu, X., Zhang, Y., Tan, Y., Wu, D., Shi, M., & Huang, H. (2023). The effect of short-form video addiction on undergraduates' academic procrastination: A moderated mediation model. *Frontiers in Psychology*, *14*, 1298361.

Taichi Inohara*
College of Media Arts, Science and Technology,
University of Tsukuba, Tsukuba, Japan

*Corresponding author: inohara@iplab.cs.tsukuba.ac.jp