Investigating User-perceived Impacts of Contextual Factors on Opportune Moments

YU-JEN LEE, National Yang Ming Chiao Tung University, Taiwan
MENG-HSIN WU, Carnegie Mellon University, United States
CHUNG-CHIAO CHANG, National Yang Ming Chiao Tung University, Taiwan
XI-JING CHANG, National Yang Ming Chiao Tung University, Taiwan
YUNG-JU CHANG*, National Yang Ming Chiao Tung University, Taiwan

In this exploratory experience sampling method (ESM) research, we examined the perceptions of 74 smartphone users regarding the opportuneness of moments for proceeding through a four-stage notification-response process: the phone generating an alert (Alert), the user roughly glancing at the notification (Glance), engaging with it (Engage), and acting on it (Act). We investigated how the moments perceived as opportune for each of the four stages related to users' self-reported values of 20 contextual factors, and how these factors influenced users' perceived opportuneness of the moments for each stage. Our results reveal that Alert and Glance stages were perceived as more distinct, with Alert being influenced by social-environmental related factors and Glance characterized by a lower threshold for what constitutes an opportune moment. The final two stages – Engage and Act – were the most similar to each other. The findings also indicated how the influence of contextual factors on perceived opportuneness of the moments varied across factors, notification types, stages, and how such variation was manifested in the likelihood, valence, and magnitude of their overall influence.

CCS Concepts: • **Human-centered computing → Smartphones**.

Additional Key Words and Phrases: Mobile notifications; mobile receptivity; opportune moment; interruptibility; ESM

ACM Reference Format:
Yu-Jen Lee, Meng-Hsin Wu, Chung-Chiao Chang, Xi-Jing Chang, and Yung-Ju Chang. 2024. Investigating User-perceived Impacts of Contextual Factors on Opportune Moments. *Proc. ACM Hum.-Comput. Interact.* 8, MHCI, Article 269 (September 2024), 28 pages. https://doi.org/10.1145/3676514

1 Introduction

In the decade since smartphones became an indispensable part of our daily lives, research on mobile notifications has proliferated. Though convenient in many ways, increasing numbers of notifications carry undesirable information that can cause phone users stress [104] and distract their attention from their current tasks [48]. Researchers have estimated that only 6.33% of notifications are considered both very important and urgent [99], with most deemed irrelevant to smartphone

*Corresponding author.

Authors' Contact Information: Yu-Jen Lee, daniel08099@gmail.com, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, Taiwan; Meng-Hsin Wu, menghsiw@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States; Chung-Chiao Chang, sean831004.cs07@nycu.edu.tw, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, Taiwan; Xi-Jing Chang, siliconcrystal.c@nycu.edu.tw, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, Taiwan; Yung-Ju Chang, armuro@nycu.edu.tw, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, Taiwan.

ACM 2573-0142/2024/9-ART269
https://doi.org/10.1145/3676514

users' needs, annoying, and disruptive [60]. The sense of notifications' disruptiveness is rooted primarily in their arrival at inopportune moments [74], such as during face-to-face social events [69] or engagement in tasks [63]. Interestingly, however, the sources of notifications have also been linked to perceptions of disruption and the opportuneness of the moments when they arrive [50], which in turn affects how fast they are attended to [12].

To reduce disruption and increase smartphone users' receptivity to delivered notifications, recent research has sought to identify opportune moments, or interruptible moments, for phone delivery of notifications about various types of content, such as reading material [24, 74], games [74], ads [74, 100], questionnaires [33, 74], interventions [49, 53, 82], and micro-tasks [13, 17, 18]. To be considered receptive, a user is typically expected not only to attend to, but also to be willing to respond to, the delivered content [33, 74]: by filling out the questionnaire, replying to the message, executing the assigned task, and so on.

Nonetheless, substantial empirical data supports the notion that the act of addressing a notification encompasses various stages that frequently occur independently or warrant individual consideration. For instance, disparities between attending to and responding to a notification have been identified in the literature. For example, Chang et al. [14] reported that alert modality was associated with users' attentiveness to notifications, but not their responsiveness. Lee et al. [50] found that attentiveness, responsiveness, and perception of opportune moments for notification delivery were all associated with different relational and contextual factors. Turner et al. [91] further divided attending into two stages, focusing and reading, whose predictive models were slightly different. In addition to distinguishing between attending and responding to notifications, other researchers have posited that the act of perceiving alerts is separate from attending to notifications [12, 89].

Given the existing literature that scrutinizes individual phases of smartphone notification engagement and discerns unique factors related to each stage [12, 14, 50, 89, 91], we hypothesize that users' perceptions of opportune moments for these discrete stages may also manifest distinct variations. For example, a user in a serene study environment may regard replying to a message as socially acceptable, but not the audible emission of an alert from their device. In contrast, the same individual may deem it ill-timed to respond to a notification in a specific social context, while permitting a cursory examination of its content.

Despite these observations, there remains limited systematic investigations into the nature of opportune moments for specific stages of notification interaction, and how these moments differ across stages. We deem that the distinctions between opportune moments for each stage may be gleaned from examining the contextual factors that render a moment favorable or unfavorable for one stage compared to others. Thus, in this exploratory study, our main goal is to provide such an examination. To achieve this, our study constructs a four-stage model—consisting of the generation of an alert (Alert), a cursory glance by the user (Glance), user engagement (Engage), and user action (Act)—inspired by and synthesized from existing research on the response process [91]. Building on this model as our framework, our aim is to explore the differences in opportune moments for each of these stages by addressing the following two key research questions:

RQ1: In what manner do the opportune moments for each stage along a notification-interaction continuum resemble or diverge from one another?

RQ2: How are the opportune moments in each stage perceived to be influenced by various contextual factors?

To answer these two research questions, we conducted an experience sampling method (ESM) study of 74 smartphone users, who provided their in-situ perceptions of opportune moments for each of the four stages of the response process for each sampled notification, as well as their perceptions of 20 contextual factors and those factors' impact on such opportuneness. This paper makes four main contributions to the literature:

- First, it shows that users' perceived opportune moments for the Alert and Glance stages are relatively distinct from the other stages, whereas the perceived opportune moments for the final two stages, Engage and Act, are more closely aligned.
- Second, it presents cross-factor and cross-stage comparisons of contextual factors' impacts on the perceived opportuneness of moments for the four stages, enabling us to identify overall trends in contextual factors' impacts as well as the contextual factors that were most and least influential on opportune moments.
- Third, it reveals how contextual factors vary across factors, notification types, stages, and how such variation is manifested in the likelihood, status, and magnitude of their overall influence.
- Fourth, it provides qualitative insights into how our participants perceived the four stages and their respective opportune moments differently.

2 Related work

To address our research question, we delved into previous studies concerning opportune moments. Through this exploration, we discerned various factors and dimensions that shape perceptions of opportune moments. Importantly, while some studies directly use the term "opportune moment", others approach the concept indirectly through related terms like interruptibility, receptivity, breakpoint, and attentiveness, among others. Additionally, many of these works explored opportune moments across different interaction stages with notifications. By bringing together these results, our aim is to foster a broader and more thorough understanding of factors related to opportune moments across different interaction stages with notifications.

2.1 Opportune Moments for Mobile Interruption

The interruptions of mobile notifications and users' receptivity to them have been extensively studied in recent years [18, 34, 57, 58, 90, 105]. Prior studies have shown that pushing notifications at an improper time reduces users' productivity [5] and increases their level of anxiety [75]. Prior analysis of the interruptions' impacts placed most emphasis on two directions: the attributes of the interruption itself, and the situations in which users are interrupted. In terms of the attributes of interruption, prior studies have revealed that users' receptivity to processing interruption is positively influenced by their perceptions of the interruption's content relevance, as well as their interest in the interruption content [31, 38, 81]. Weber et al. [99] discussed influencing factors, such as notification urgency and notification importance, of perceived notification interruption. In addition to content, Lee et al.'s work [50] indicated the influence of interruption source, such as the sender-recipient relationship, on users' receptivity for and their perceived disturbance of incoming message notifications. In terms of interruption situations, existing studies have identified contextual factors in desktop settings [6, 59, 90] and mobile settings [47, 50, 63] that influence the perception of interruptions. For example, Mehrotra et al. [64] suggested that individuals' receptivity to notification is influenced by their emotional states, and they attempted to estimate the opportune moment for delivering certain types of content. Another work from Mehrotra et al. [63] revealed that individuals' perceived disruptiveness of a notification can be influenced by the complexity of the on-hand task that the individual is currently engaging in. Additionally, Cha et al. [10] found that personal contextual factors together with contextual factors associated with everyday home routines influence users' opportune moments for proactive conversational interactions. Pejovic et al. [72] found that receptivity of messaging notifications is negatively correlated with individuals' level of activity engagement. Chang et al. [102], on the other hand, suggested that users' preferences for when to read notifications are contingent upon their motivations for doing so. They identified four reasons for which individuals opt to read notifications during an activity rather than before or after it. Thus, these studies indicate that perceiving opportune moments can be quite complex and involve various factors. They provide us with valuable insights into comprehending the relationship of the factors and the degree to which they may influence users' perceptions of opportune moments.

Another line of research in interruptibility leverages machine learning techniques to model and predict interruptible and opportune moments (e.g. [15, 61, 66–68, 71, 74, 79, 88, 92]). For example, Pielot et al. [74] built a machine-learning model that predicts opportune moments for proactive user engagement with the notification, or clicking and subsequently engaging with the notification content. Okoshi et al. [68] estimated interruptibility and adaptive notification scheduling. Visuri et al. [97] built a machine-learning classifier based on both notification content and contextual sensing to predict whether the user wants to see a new notification or not at the moment. Hudson et al. [36] employed machine learning algorithms to construct a predictive model of human interruptibility with sensors. Iqbal and Bailey [37] computed the cost of interruption with high accuracy, allowing future systems to make more effective decisions about when to interrupt. Khan et al. [44] devised the machine learning rule-based methodology to predict user behavior-based rules based on providing context on smartphone notification services. Forlivesi et al. [32] developed a multi-modal wearable-only system to model interruptibility. Chen et al.[15] predicted the opportune times to deliver notification in virtual reality. Albeit not using interuptibility, Pielot et al. [77] predicted users' boredom from their mobile phone usage, and found that people are more likely to engage with recommended content when they feel bored. Similarly, Chen et al. [16] predicted smartphone users' killing-time behaviors based on screenshot and sensor data. Isaacs et al. [40], on the other hand, introduced the idea of microwaiting to describe brief intervals of idleness during an ongoing activity, and reported that people frequently request content during such moments. Dingerler et al. [24], built a classifier to detect the opportune moment for reading. Van et al. Draxler et al. [26] proposed both schedule-based and activity-based triggers for microlearning, which contribute to enhancing learners' engagement. Mendel et al. [65] developed regression models that predict moments when smartphone users need support, highlighting breakpoints as particularly opportune times for notification delivery. Additionally, Van et al. [96] identified smartphone sessions and demonstrated that breakpoints in these activities are predictable. Several other studies have also demonstrated that notifications that are pushed during breakpoints are dealt with significantly more quickly [30], with less frustration [38], and are perceived as less interruptive [69].

However, despite the extensive body of existing research on opportune moments and related concepts, there is a notable gap in analyses focusing on the influence of factors on different stages of the response process. In the following section, we will discuss previous works that offer empirical evidence supporting the differentiation between these stages.

2.2 Stages in the Notification Response Process

Recognizing that smartphone users go through multiple stages to complete a responding process has been explicitly argued, or implicitly indicated, in prior notification and interruptibility research. Explicitly arguing for decomposing the responding process into multiple stages, Turner et al.'s framework [91] proposed four stages: react (gain the user's attention through sound, vibration), focus (after choosing to react, e.g., turning the screen on), read (read a fuller message by consuming the notification and entering the relevant application), act (act on the content, e.g., send documents in reply to the email). Later, Turner et al. [92] distinguished reachability (indicating whether a response will at least be started, or not) [91], engage-ability (indicating whether a response will be started but abandoned without consuming the notification) [91], and receptivity (indicating whether the user is receptive to the notification content and consumes it) [31, 61, 91] and developed models for predicting users' theses statuses. In a recent study by Tseng et al. [29], the researchers introduced a comparable set of notification-interaction phases, encompassing Notice, Glance, Read, and Act. They posited that an individual's decision regarding the appropriate device (e.g., smartphone, smartwatch, tablet, or computer) for advancing to a particular interaction stage for a notification is contingent upon the device's suitability in terms of its status and characteristics for that specific interaction.

In addition to these studies, numerous research efforts have focused on users' specific behaviors related to "attending" to notifications. The majority of this research explores how quickly users click notifications upon their arrival and how frequently they attend to them, clearly distinguishing these behaviors from responding to notifications [22, 51, 64, 75, 76]. For instance, Dingler et al. [22] used phone logs to quantify user attentiveness and found that users were attentive to messaging notifications for approximately 12 hours per day. Pielot et al. [76] built machine learning models to predict users' attentiveness to instant messages. Mehrotra et al. [64] computed three time-based metrics—average seen time, average decision time, and average response time—to record users' attentiveness and receptivity to notifications across each dimension. Conversely, other researchers have focused on predicting responsiveness to phone notifications rather than attentiveness [46, 50].

However, prior research has shown mixed results regarding the distinction between attentiveness and responsiveness. [12, 14, 19, 23, 50, 101]. Specifically, Chang et al. [14] found that the ringer mode of a notification affects users' attentiveness to message notifications but does not impact their responsiveness. Similarly, Lee et al. [50] identified sender relationship closeness as an indicator of attentiveness but not of responsiveness in instant message notifications. In a recent study, Lee et al. [52] revealed that responding to a message involves a complex decision-making process, incorporating several factors and sometimes strategic considerations, thereby distinguishing it from merely reading a message. On the other hand, recent qualitative studies have begun to reveal that the distinction between attentiveness and responsiveness is not perceived uniformly by all users. For instance, Wu et al. [101] reported that while some users consider the timing for responding and attending to be identical—often responding to an instant message immediately after reading it—others perceive attending as merely noticing the presence of a message, thereby distinguishing the timings for these actions. Furthermore, Chou et al. [19], in their study on users' perceptions and practices around read-receipts in instant messaging services, revealed that users sometimes intentionally defer reading a message until they are ready to respond, as reading without responding may be viewed as socially inappropriate.

Given these mixed results and the varied interpretations and practices surrounding the distinction between these stages, it remains unclear how the opportune moments for these two stages should be considered similar or different from each other. Thus far, there is limited understanding about how users perceive opportune moments for different stages and if distinct factors influence each stage differently. Based on the literature, we outlined all substantial contextual factors and organized them into five different contexts in this work, shown as the following: personal [1, 10, 47, 48, 64, 77, 83, 105], activity [28, 39, 61, 62, 64, 71, 72, 80, 83], social environment [27, 28, 43, 62, 70, 89], notification content [31, 55, 63, 81, 97], and notification sender [3, 4, 50, 63, 85, 98, 105]. Beyond the studies which consider individual associations between specific stages and factors, this paper provides a more comprehensive investigation into how the aforementioned factors are associated with the different stages throughout the responding process.

3 Framework Design

To address the research questions, we referred to Turner's framework to decompose the notification process into four stages. We then gathered 20 factors from the aforementioned related studies and categorized them into four groups.

3.1 The Four Stages of the Notification-response Process

Based on our literature review, we decomposed the notification-response process into the following four stages:

(1) `Alert`: The device generates an alert to attract the user's attention either via vibration, sound, or screen light when the smartphone is away from participants, or pop-up notifications while they are actively using the device.
(2) `Glance` [7]: The user shifts his/her attention to the notification and quickly skims it to gain a rough idea of its purpose, either as soon as it arrives or when browsing multiple notifications in a notification drawer. At this stage, the user does not engage in any kinds of hand movement (e.g., tap, click) to interact with the notification.
(3) `Engage`: This stage occurs when the user chooses to tap the notification to read its full content. This stage involves the user's hand movement.
(4) `Act`: If the notification requests or implies a request for action beyond merely viewing the notification (e.g., replying to an IM message or performing a requested task), and the user agrees that such action is necessary, s/he acts upon the notified content.

The naming and definition of these four stages have been slightly modified from Turner et al.'s [91] React, Focus, Read, and Act stages, to make them better align with our research aim. For example, Turner et al.'s first stage, *react*, focused on the user's status; whereas we adopted the name Alert based on other research showing that the perceived opportuneness of a moment for a phone to generate an alert may involve perceived social appropriateness [14, 69, 89] rather than whether doing so is physically possible. Glance, meanwhile, reflects the characteristics of a common notification-attending behavior [64]. On Android systems, in this stage, users skim through a notification title and its short description, instead of taping into it to see its full content (unless it is unusually short). Another reason we did not adopt the terms "Focus" and "Read" was that our pilot participants were commonly confused between them. An additional reason that we adopted the term Engage from Turner [92] in place of "Read" for the third response stage was that we felt Engage better characterized an action (i.e., physically tapping into the notification) performed to access a notification's full content, distinct from merely glancing at it. We emphasized this distinction between Glance and Engage to all the participants, and made sure they understood the distinction, before they started the study. Finally, we adopted Turner et al's [91] original term Act, but emphasized to the participants that this stage implied doing something more than simply reading the full notification content. Common examples of Act include responding to the message, performing tasks related to the notification, etc.

3.2 Contextual Factors

We investigated the relationship to each response stage to 20 contextual factors, classified into five categories. These factors were selected because all had previously been found to influence users' receptivity to interruptions and incoming notifications in some way.

- `Activity Context`: These factors describe the user's perceived characteristics of the users' current task at hand, including task importance [86], task complexity [63], engagement level [31], and busyness level [63].

- `Personal Context`: These factors describe the user's perceived their own psychological status, including happiness [64], stress [64], boredom [77], fatigue [64], and loneliness[75].
- `Social Environmental Context`: These factors describe the user's perceived characteristics of the user's current environment, including privacy [89], social norms [9], number of people the user is interacting with [89].
- `Notification Content Context`: These factors describe the user's perceived characteristics of the notification content, including urgency [99], importance [99], attractiveness [8], relevance with the task at hand [31], effort required to act on notification content [63].
- `Notification Sender Context`: These factors describe the user's perceived relationship characteristics with the sender of the communication notification (e.g. instant messaging, e-mail, and social media), including importance [50], closeness [50], relationship type [63].

4 Methodology

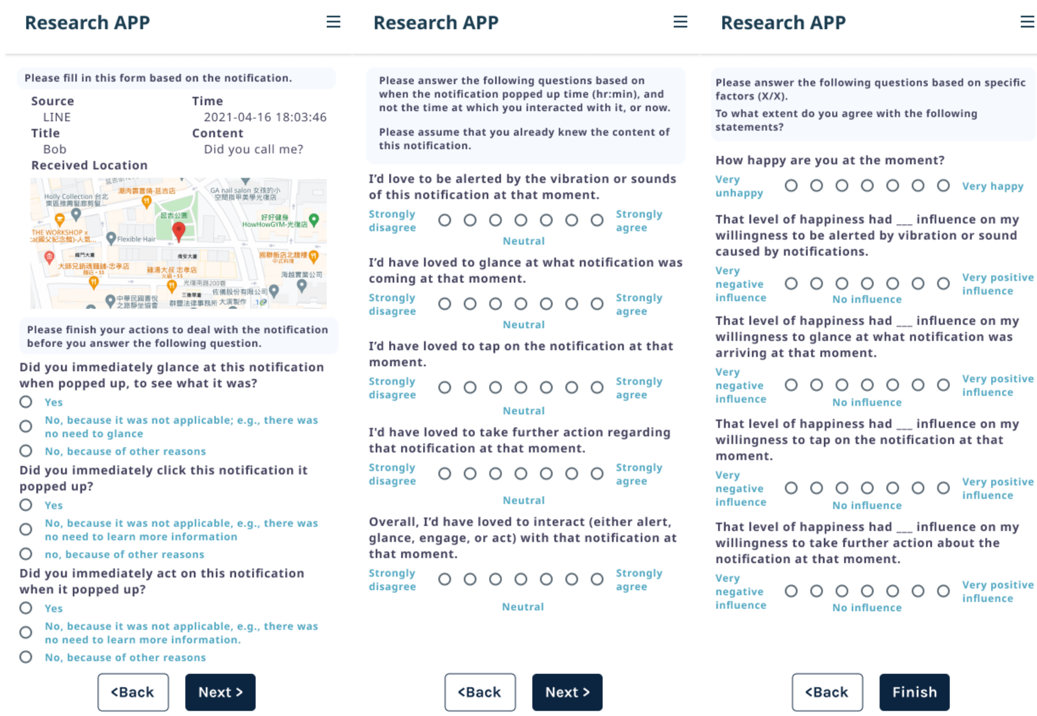
We conducted a study with 74 smartphone users via an Android app that monitored their phone status, sampled their notifications, and sent ESM questionnaires that asked them to report their perceptions of the opportuneness of the moments when the sampled notifications arrived. They also reported their perceptions of the intensity of 20 contextual factors and those factors' impact on such perceived opportuneness. All design decisions were made after conducting a pilot study with 10 participants to refine the ESM mechanism and wording used in the study. Details of the ESM study are provided below.

4.1 Experience Sampling Study

We developed a research app that recorded all notifications on the phone, sample specific notifications, and delivered ESM questionnaires [93, 95] for these sampled notifications. We explain these in more details in the following subsections.

4.1.1 ESM Questionnaire. An ESM questionnaire asked the participants to report what stages of the response process they have done and their perceptions of the opportuneness of the moments the sampled notification had arrived, for each response stage and overall. Specifically, we first asked participants whether they had glanced at, clicked on, or acted on the notification when it arrived, and reported if the notification was not applicable to that action, as illustrated in Figure 1a. Then, we assessed the opportuneness of such moments in terms of the participant's *willingness to proceed to a given stage*, on a seven-point Likert scale: with an ESM response that s/he *"would love to"* proceed to that stage equating to an opportune moment for that stage, as shown in Figure 1b. We adopted the notion of *willingness* because it was both specific enough for our research purposes and intuitive enough for our pilot participants to answer confidently, and because it has previously been deemed a crucial element of receptivity [31].

Next, the participants answered questions about their perceptions of each of the 20 contextual factors, as shown in Figure 1c. Specifically, they were first asked about their perceptions of the value of a factor (e.g., "How happy are you at the moment?"), again on a seven-point Likert scale. Then, they were asked to rate their perceptions of the impact of that factor on their willingness to proceed to each of the four stages. In the tutorial video of study instructions, we emphasized that participants should rate the factor value and its impact based on their most intuitive judgement at the exact moment when their devices received the sampled notification. We included both questions asking the value as well as the impact of a factor because we recognized that the relationship between the value and the impact of a factor on users' perceived opportune moment for a particular stage can be complex. For example, it is possible that a notification is considered important and meanwhile the user rates its high importance as either having a positive or a negative impact on



(a) (b) (c)

Fig. 1. The research application includes questions that ask: (a) the stage to which the participant proceeds of the sampled notification; (b) the extent to which the participant is willing to proceed to each stage; (c) the perceived value of each contextual factor and its impact on the participant's willingness to proceed to each stage.

her decisions to read the notification, depending on whether she wants to defer it to a later point or not given her availability in the moment [54]. By including both questions asking the value as well as the impact, we were able to obtain a more in-depth understanding of how the value of the factor relates to the impact of the factor on the user's perceived opportuneness of the moment.

Therefore, there were five questions for each factor (a value plus its impact on the four stages). The impact-related questions were on a seven-item scale with a "no influence" option in the middle, and the three items on the left-hand and right-hand sides indicating positive and negative impact, respectively.

Given the large quantity of ESM questions, and the above-mentioned drawback of lengthy ESM questionnaires, we divided the questionnaire into two sections. This division enabled participants to respond to only one-third of the questionnaire when they received it, and then complete the remaining parts at other times. Specifically, the questions in the first section included 1) the response stage to which the participant had proceeded with a particular sampled notification; 2) their willingness to go through all four response stages with it; 3) factors that scholars consider more difficult to recall later, including personal context factors [42]; and 4) situational details that would aid in recalling the context of the sampled moment during the second phase of the questionnaire. These details included their ongoing tasks, the number of people they were interacting with, and the identities of the notification senders. We incorporated these elements because previous research has highlighted that in-situ information is often challenging to recall, potentially leading to inaccuracies in ESM responses [94]. Inspired by the Day Reconstruction Method (DRM) [21, 25, 41, 45, 56], which involves providing participants with contextual information to help them reconstruct their days and recall daily-life details, we included these situational details in the second part of the questionnaire.

Through our pilot study, we observed that participants typically completed the first phase in two minutes and the second phase in six minutes. After finishing phase one, the participants had up to 12 hours to finish phase two, which contained 13 factors (or 10, if the sampled notification did not include any sender information), all of which were considered more recallable by the pilot participants than the seven factors always included in the first phase. These 13 factors were drawn from all factor categories other than "Personal."

4.1.2 ESM Mechanism. ESM questionnaires were delivered by an Android app we developed. After the participants installed the research app, they first were prompted to select 14-hour windows during which they were willing to receive ESM questionnaires each day, which could be different for weekends and weekdays. The research app sampled notifications based on their scarcity: i.e., notifications from less-sampled apps received higher weightings in sampling. This decision was made because smartphone users received more certain type of notifications than others [78, 81, 87]; thus, we deemed it important to capture various types of notifications to prevent our ESM responses being strongly biased due to imbalances in notifications' popularity. The research app also considered the time that had elapsed time since the start of the 14-hour daily window, and the time when the previous ESM questionnaire had been completed. This allowed us to balance our need for the app to wait for rare notifications against the possibility that a whole day might pass without any rare notifications being received. That is, the app was more heavily weighted toward sampling rare notifications near the beginning of the window, but the more common notifications were increasingly weighted as the end of the window drew nearer. Our weighting parameters and elapsed-time thresholds were iteratively tested with pilot participants, until our app could capture notifications from various sources in relatively balanced, yet sufficient numbers. Whenever an arriving notification was sampled, the research app waited for one minute, in case the participant immediately reacted to it. Then, the research app issued an ESM questionnaire through a phone notification that did not generate any sound or vibration alert, to minimize the disruption it might cause. The ESM expiry threshold was set to 30 minutes. In our pilot study, we validated that this expiry threshold for our ESM questionnaire was sufficient to capture participants' low-availability moments—thus not overly biased towards moments of high availability for notifications—as also indicated by prior ESM studies (e.g., [50, 55]). Additionally, this duration was short enough to ensure participants could clearly recall their experiences. Once a participant completed phase one of an ESM questionnaire, they would not receive another for at least one hour, and the app was programmed to issue no more than eight questionnaires per participant per day.

4.2 Study Procedure

Once a participant signed up for the study, they received a consent form and a short tutorial video via email detailing the study's scope. We guided participants through the consent form, explaining the research app's functionality and the nature of the data it collected and stored. Participants were informed that their data would be anonymized, with all personally identifiable information removed, and stored on our secure server accessible only to the research team. They were also informed of their right to withdraw from the study at any point without any consequences. Upon agreeing to participate and signing the consent form, participants were redirected to our research web page, which included a link to download the research app and a tutorial video on how to install and configure it. .

The research app started delivering ESM questionnaires on the day following successful installation, and continued doing so for 14 days, unless there were any days on which the participant did not complete at least one questionnaire. In those cases, the study period was extended by one day for each day that was missed, up to a maximum of seven extra days or 21 days in all, to prevent the experiment from being too long. To enrich our data qualitatively, we conducted a 45-minute in-depth semi-structured interview with 30 of the participants whom we selected either because 1) their patterns of ESM responses were distinctive from those of the other participants, or 2) whose ESM responses varied markedly from one questionnaire to another. Thus, our interviewees were relatively diverse. The purpose of the interviews was to understand how the participants perceived the similarity and differences between the opportune moments for the four stages. The participants received NT$20 (approximately US$0.72) for each completed ESM questionnaire, and an additional NT$300 (approximately US$11) if they took part in an interview.

4.3 Recruitment and Participants

We posted our recruiting message in a number of forums and Facebook pages, including some pages intended specifically for the recruitment of research subjects, and others for general use by residents of specific Taiwanese cities. The recruitment posts led to the sign-up form. During the recruitment process, we aimed to balance the participants' genders, ages, and occupational backgrounds. A total of 78 participants joined the study and 74 completed it. Of the four that did not complete it, one withdrew because the research app did not function properly, and the other three simply ceased participating and never contacted us again. Among the remaining 74 participants, 43 were females and 31 were males. They ranged in age from 20 to 55 (M=25.6), and 44 were students. The remainder were from a variety of job sectors, including manufacturing, information technology, and entertainment, among others. Due to extensions for skipped days, as explained above, the mean duration of participation was 15.42 days (SD=2.08), not including the four participants who withdrew. Of the 30 participants who participated in post-study interviews, 17 were female and 13 were male, and 10 were students. The mean age of the interviewees was 28.5 (Max=55, Min=20, SD=9.40).

4.4 Data Cleaning and Analysis

During the study period, the participants collectively received 326,560 phone notifications and completed or partially completed 5,451 ESM questionnaires. 371 questionnaires were incomplete and excluded from further analysis. Also, we ignored the first three ESM questionnaires completed by each participant, on the assumption that their initial unfamiliarity with the questions might skew their early responses. This led to the removal of a further 207 sets of questionnaire answers, resulting in a final dataset of 4,873 ESM responses. On average, each participant completed 65.8 of the questionnaires that were analyzed (Max=113, Min=3, SD=32.58).

Note that our ESM asked the participants whether they had proceeded to each of the response stages (except Alert, since alerts are system-initiated), and allowed them to tell us if any such stage was not applicable to the sampled notification. For analysis of a specific stage where this "not applicable" answer option was selected, that ESM questionnaire was not included in our analysis of that response stage. As a result, the numbers of ESM questionnaires that were included in all quantitative analyses were 4,873 for Alert, 3,883 for Glance, 3,189 for Engage, and 3,028 for Act.

For statistical analysis, we employed a mixed-effects linear regression on dependent variables derived from ratings, and mixed-effects logistic regression on binary dependent variables. We chose mixed-effects models for two primary reasons. First, each participant contributed repeated

General	0.735	0.889	0.962	0.962
	Alert	0.745	0.690	0.687
		Glance	0.894	0.867
			Engage	0.965
				Act

Fig. 2. Spearman's rank correlation coefficients (ρ) of five types of opportune moments: General, Alert, Glance, Engage, and Act. All correlations are significant at the 0.001 level.

observations over the study period. This approach allows for accounting for individual differences, enabling us to estimate fixed effects on the dependent variables while controlling for random effects, with participant IDs used to account for individual variance. Second, the robustness of mixed-effects regression against violations of the normality assumption has been demonstrated in prior research [84], validating its appropriateness for our dataset.

We used affinity diagramming [35, 103] to analyze our qualitative data, which involves iterative labelling and grouping of notes that were transcribed interview audio recordings. The research team collaboratively sorted each note and went through the entire affinity diagram together in each labelling and grouping session. The process led to 3 high-level themes: 1) opportune moments for each of the four stages, and 2) influence of contextual factors on each of the four stages, and 3) the difference between the four stages. The highlights of our qualitative findings are presented in the following sections.

5 Results

5.1 Correlation between the Four Stages

We began by calculating the correlation between participants' ratings for the perceived opportune moments for notifications across four distinct stages: Alert, Glance, Engage, and Act, as well as their ratings for general opportune moments. We used Spearman's rank correlation coefficients to assess these relationships, with all correlations presented in Figure 2 being significant at the 0.001 level. The results revealed a strong correlation among the stages, with coefficients all above 0.687. The Engage and Act stages showed the highest correlation with the general opportune moments (0.962). These stages also showed a particularly strong mutual correlation (0.965). In contrast, the Alert stage was markedly distinct from the other stages, exhibiting the three lowest pairwise coefficients: 0.745 with Glance, 0.690 with Engage, and 0.687 with Act. This stage also had the weakest correlation with the general opportune moments, as indicated by a coefficient of 0.735. These initial results collectively suggest that participants' perceptions of opportune moments for notifications align more closely with times they find suitable for fully engaging with the content and taking action than with times for merely being alerted to its presence.

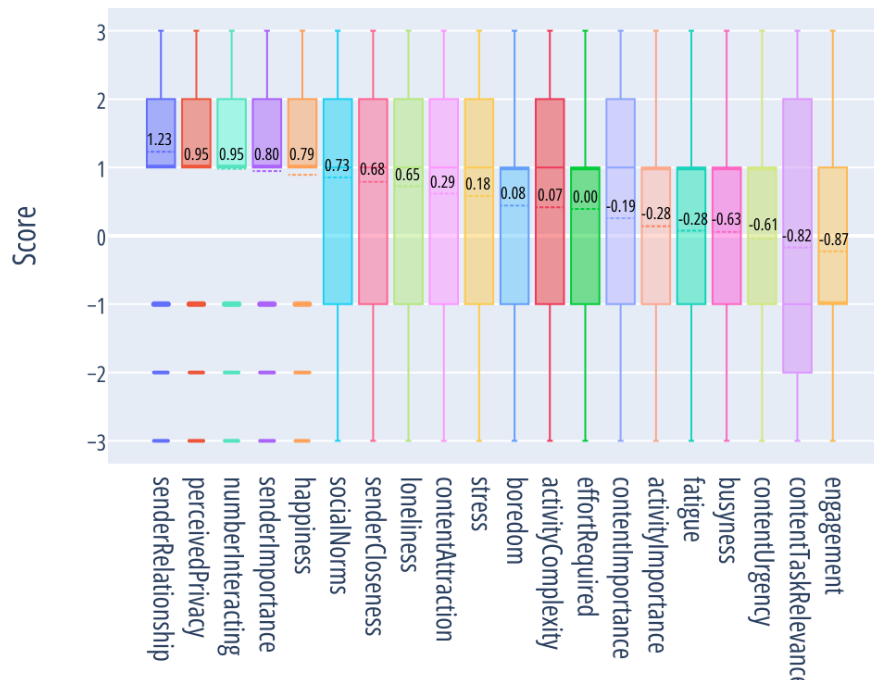


Fig. 3. The boxplot displays the distribution of participants' perceived influence scores for the 20 contextual factors on the opportune moments for notifications, averaged across the four stages. Dotted lines within each box plot represent the mean influence scores. The box plots are organized in descending order, ranging from the most positive to the most negative influence scores, from left to right.

5.2 Perceived Influences of the Contextual Factors on Opportune Moments

As mentioned earlier, the participants rated the impact of various contextual factors on the opportuneness of advancing through each response stage. They used a seven-point Likert scale for this rating, which ranged from -3.0 to 3.0. On this scale, a response of 0 indicated "No influence." To analyze these ratings, we defined and examined two metrics: *influence likelihood* and *influence score*.

- Influence Score: the mean value, between -3.0 and 3.0, of all participants' non-zero ratings of the influence of that factor on the opportuneness of the moment for a given stage. In other words, this score considers a factor's influence only if it was perceived as having either a positive or a negative influence on the opportuneness of the moment.
- Influence Likelihood: a value between 0% and 100% reflecting the likelihood of a contextual factor being perceived as having an influence on the opportuneness of the moment; i.e., of it being given a non-zero rating, regardless of whether that rating was positive or negative.

To compare influence scores and likelihoods among the contextual factors and stages, we conducted pairwise comparisons of all 20 contextual factors within our regression models. In these

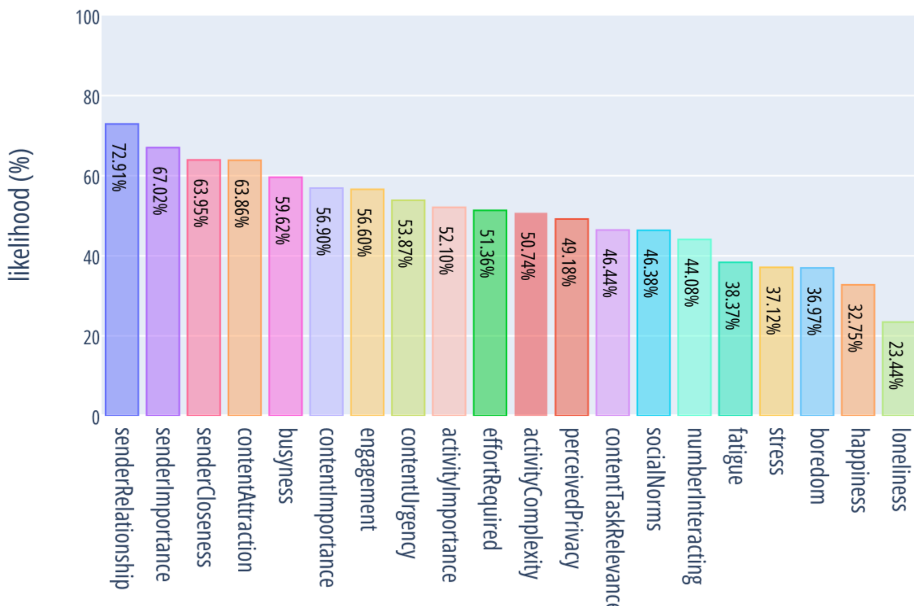


Fig. 4. The bar chart illustrates the influence likelihood for the 20 contextual factors on the opportune moments for notifications, averaged across the four stages The bars are arranged in descending order, from the highest to the lowest likelihood, moving from left to right.

models, each contextual factor was included as an independent variable. Influence score (numeric) and influence likelihood (binary) were included as dependent variables. Given the multiple comparisons were conducted across factors and stages, we applied the Bonferroni correction to adjust our significance levels [2]. For these tests, a stringent significance level ($\alpha = 0.05/n$) was established, where 'n' represents the number of comparisons specific to each factor or stage in question.

We present the results of the influence scores shown in a box plot, depicting the mean and distribution in Figure 3, and influence likelihoods presented in a bar chart in Figure 4. A comprehensive visualization of the influence score and likelihood of all contextual factors is provided in Appendix in the supplementary file.

5.2.1 The Overall Influence of Contextual Factors on Opportune Moment. Sender relationship emerges as the predominant contextual factor, with a positive influence score of 1.23, which is statistically significantly higher than those of other factors (all $p < 0.001$). Following closely, Sender Importance ranks second with a score of 0.95, also significantly exceeding comparisons (all $p < 0.001$). Conversely, Content-Task Relevance and Content Urgency exhibit the most substantial negative influences, with scores of -0.87 and -0.82 respectively, without significant differences between them ($t(15895) = 1.14$, $p = 0.253$). Several factors were consistently associated with low influence scores, both positive and negative. Specifically, factors such as Boredom (0.00), Activity Complexity (0.07), Required Effort (0.08), and loneliness (0.18) were perceived as minimally influential on average.

In terms of influence likelihood, factors related to the sender, again, were most frequently perceived as influential. Specifically, Sender Relationship was perceived as influential 72.91% of the

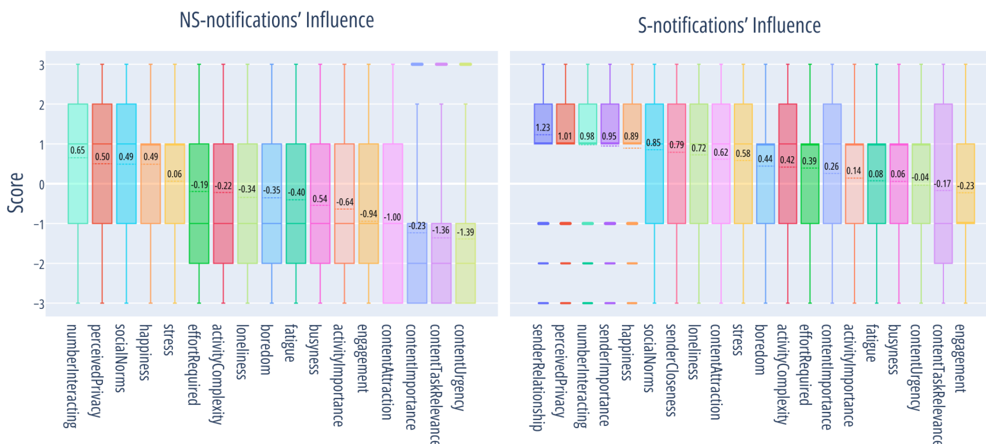


Fig. 5. The boxplot displays the distribution of participants' perceived influence scores for the 20 contextual factors on the opportune moments for notifications without involving a sender (NS-notifications) on the left and involving a sender (S-Notifications) on the right. Dotted lines within each box plot represent the mean influence scores. The box plots are organized in descending order, ranging from the most positive to the most negative influence scores, from left to right.

time, followed by Sender Importance (67.02%), and Sender Closeness (63.95%). Among non-sender factors, Content Attraction was most frequently noted as influential, 63.93% of the time. In stark contrast, Loneliness was least likely to be perceived as having an influence, noted by only 21.1% of the time, and its likelihood was statistically significantly lower than that for other factors (all $p <$ 0.001).

Combining the two sets of results reveals varying influences of certain factors. For instance, while the influence of Content Attraction was prevalent, it exhibited a relatively low average influence score. This suggests that although it was often perceived as playing a role, its impact was considered relatively minor. In contrast, the influence of factors such as Happiness, although less prevalent, was significant when perceived; its average influence score was among the highest. These results underscore that different contextual factors play distinct roles in shaping the perception of opportune moments.

5.2.2 Differences in Influential Factors between Sender and Non-sender Notifications. Given that not all notifications involve a sender, we differentiated between notifications with a sender (S-notifications) and those without (NS-notifications). S-notifications typically originate from messaging apps and social media platforms. Our findings, illustrated in Figure 5, indicate several noteworthy trends. First, influence scores for NS-notifications tended to be more negative compared to S-notifications. For instance, while loneliness averaged a positive influence score (0.72) in S-notifications, it shifted to a negative score (-0.34) in NS-notifications. Second, the four content-related factors exhibited particularly strong negative influences in NS-notifications; however, three of these factors (except for Content Task Relevance) were perceived positively in S-notifications. This notable disparity suggests that participants' perceptions of content-related factors significantly differed depending on the presence of a sender. Third, three environmental factors—Number of People Interacting, Perceived Privacy, and Social Norms—and one personal factor, Happiness, consistently rank as top positive influences across both notification types. This consistency implies

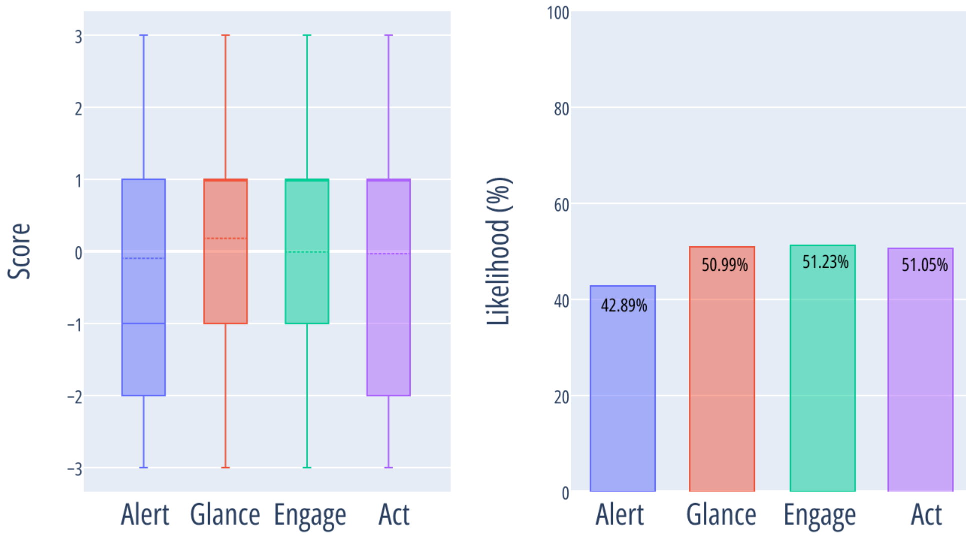


Fig. 6. The boxplot on the left displays the distribution of participants' perceived influence scores for the 20 contextual factors, organized by stage, on the opportune moments of the four stages. Dotted lines within each box plot indicate the mean influence scores. The bar chart on the right illustrates the likelihood of these 20 contextual factors being perceived as influencing the opportune moments for the four stages.

that these factors exerted a stable influence from the participants' perspective, largely independent of notification type.

5.2.3 Influence of Contextual Factors on Opportune Moments Across the Four Stages. We found that in the Glance stage, contextual factors were perceived as having a more positive influence on the opportuneness of the moment for proceeding with the Glance stage compared to other stages, as shown in Figure 6. Notably, it is also the only stage that received a positive mean influence score (Glance (M = 0.18, SEM = 0.01) vs. Alert (M = -0.10, SEM=0.01, $t(74147)$ = -21.04, p < .001); vs. Engage stage (M = -0.07, SEM = 0.01, $t(66269)$ = -12.27, p < .001; vs. Act stage (M = -0.03, SEM = 0.01, $t(64723)$ = -14.33, p < .001). In other words, participants tended to deem a given situation as opportune for glancing at notifications but inopportune for the other stages. In contrast, the Alert stage was on average perceived as the most negatively influenced by contextual factors among the four stages (vs. Engage stage (M = -0.07, SEM = 0.01, $t(68023)$ = 6.44, p < .001; vs. Act stage (M = -0.03, SEM = 0.01, $t(66477)$ = 4.38, p < .001). Notably, despite a slight decrease in influence scores in the latter two stages, no significant difference was observed between them ($t(58601)$ = -2.44, p = 0.115), indicating a similar perception regarding opportune moment across these stages.

In addition to the general trend, we found that specific contextual factors exhibited notably stronger influences on particular stages when these factors were at a certain status. For example, we observed that activity-related factors (e.g., busyness) in a particularly strong negative influence

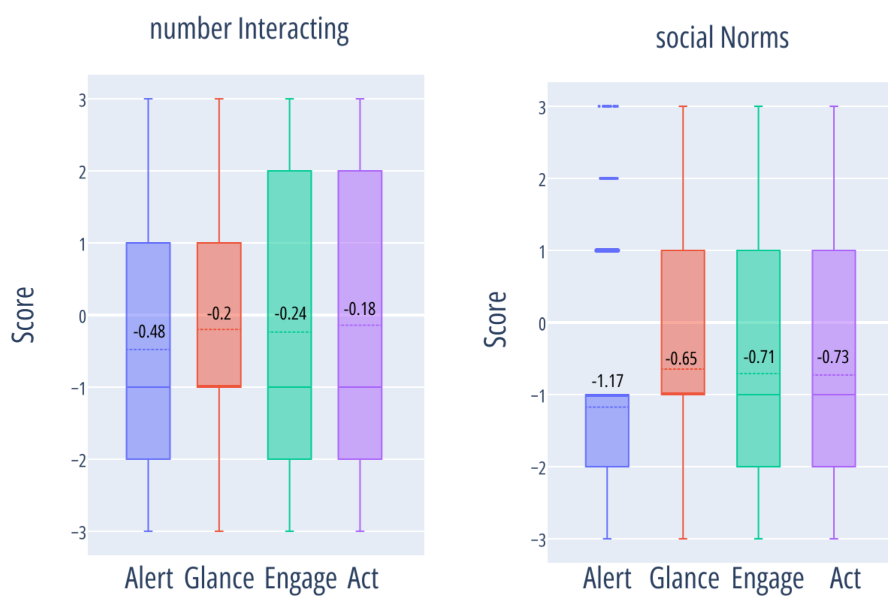


Fig. 7. The boxplot on the left displays the distribution of participants' perceived influence scores for the factor Number of People Interacting on the opportune moments of the four stages. The boxplot on the right shows the distribution for Social Norms under the same conditions. Dotted lines within each box plot indicate the mean influence scores.

on the Act stage when they were in high status (i.e., busy). The mean influence score was -1.29 (SEM = 0.02), more negative than those of other stages (vs. Alert stage: M = -1.20, SEM = 0.02, $t(9949)$ = -4.35, $p < .001$; vs. Glance stage: M = -1.07, SEM = 0.02, $t(9589)$ = -8.86, $p < .001$, vs. Engage stage: M = -1.25, SEM = 0.02, $t(9058)$ = -2.25, p = 0.024). Similarly, two social-environmental factors—Number of People Interacting and Social Norms—as shown in Figure 7, when in a positive status (i.e., interacting with a group of people and perceiving high social norms), also exhibited a particularly strong negative influence on the Alert stage, significantly more negative than it on the Glance stage (Social Norms: Alert stage M = -1.17, SEM = 0.02 vs. Glance stage M = -0.65, SEM = 0.05, $t(1804)$ = 5.36, $p < .001$; Number of People Interacting: Alert stage M = -0.48, SEM = 0.08 vs. Glance stage M = -0.20, SEM = 0.05, $t(342)$ = 2.05, p = .045)

In addition, we also observed differences in influence likelihood among the four stages. Specifically, the Alert stage was less likely to be perceived as influenced by contextual factors (42%) compared to the other stages (vs. Glance (M = 50%, z = 34.12, $p < .001$); vs. Engage stage (M = 51%, z = 34.43, $p < .001$); vs. Act stage (M = 51%, z = 31.89, $p < .001$)). However, the Glance, Engage, and Act stages did not show significant differences in influence likelihood. Notably, Perceived Privacy exhibited a unique characteristic; in perceived low-privacy environments, this factor were the most likely to be perceived as negatively influencing the Alert stage (52%) compared to the others (vs.

Glance (M = 45%, z = -4.89, p < .001); vs. Engage stage (M = 47%, z = -2.33, p = 0.019); vs. Act stage (M = 48%, z = -2.094, p = 0.036)).

To sum up, these findings indicate that participants tended to perceive certain situations as more or less opportune for proceeding to specific stages than to others. Furthermore, some contextual factors were perceived to have a particularly strong influence on specific stages when they were in a certain status. Our qualitative results below provide further explanations for these observations.

5.3 Self-Reported Inter-Stage Differences and Their Opportune Moments

In line with our quantitative results, our interviewees reported a variety of inter-stage differences in what they considered opportune moments, and provided various explanations for such differences. These qualitative results provide insights into what constitute opportunity moments for these stages and serves as the final element of this paper.

5.3.1 The Alert Stage Was about Drawing Attention and Disruption. The interviewees mentioned several fundamental differences among the four stages, in terms of both their outcomes and the cognitive and physical effort they required.

When talking about the Alert stage, they noted that an alert is intended to draw people's attention, which – when it happened – would also disrupt and distract them when the moment was inopportune. Therefore, they noted that when they shifted their attention to the alert, they found themselves wishing that this effort *would be worthwhile*. Key notification characteristics that the interviewees mentioned in this context included importance, urgency, relevance, and attractiveness, which resonates with the properties mentioned in prior research [31, 99] and in our quantitative results as desirable for notifications that generate alerts. As P12 noted, *"Of course I hope that the [unimportant] notifications will not generate any vibration. I'll deal with them when I'm more available."*. On the other hand, various interviewees mentioned that alerts would draw not only their own attention, but others'. Therefore, they were particularly concerned about disrupting social interactions and their immediate settings, such as classrooms (P01, P04), offices (P07, P27), libraries (P22), movie theaters (P04), public transportation (P08), and courts (P27). For instance, P04 said that the theater was *"a place that people have a consensus to not make any sounds, [...] I feel like I have to obey public consensus in the theater environment."*

As well as these attention-drawing phenomena, some participants mentioned that alerts could sound pushy and cause them stress, especially when they were already in that state. As P5 explained, *"I'm used to turning off the sounds and vibrations because I feel the notification is urging me to do something."* Finally, some were concerned about not noticing alerts when they were engaged in an ongoing task. Alerts were considered particularly opportune, or even desirable, when they were anticipating incoming communication. Unfortunately, the state of waiting for notifications was not listed as a contextual factor in our ESM.

5.3.2 The Glance Stage as Seeking Targets Worthy of Attention-switching. All participants considered Glance the stage at which they decided whether to switch their attention to obtain a quick, rough overview of the arriving notification. They were most concerned about the post-switching cost, e.g., the effort of resuming a cognition-demanding task. As P10 noted, *"It takes much effort to relocate my attention back to the complex task"*. Notifications that the interviewees considered more worth switching attention to were those described as interesting (P04, P14, P28), relevant (P09, P14), important (P28) or urgent (P07, P22), or sent by people they were close to (P09, P12, P14). They also expressed a consensus that being bored and killing time were opportune moments for glancing (P01, P02). One (P7) mentioned that glancing at new notifications helped with recovery from a negative mood: *"When I am unhappy [...] I still take a glance at messages, because it might make me feel better."*

5.3.3 Engage: Status Transformation from Transient to Committed. When talking about the Engage stage, the interviewees said they perceived its key characteristic as a transformation from the tentative and transient attention and effort required by Glance to a greater commitment of their physical effort, cognitive effort, and time. They made this commitment when they thought it essential to see a notification's full content, such as when it was urgent (P06, P21), time-sensitive (P05, P19) and relevant (P07, P10, P22). Such commitment was further extended if the content entailed more actions, and this was further influenced by relational factors. That is, the interviewees reported that they responded to messages faster when they had been sent by family members (P16, P18), supervisors (P02, P04), close friends (P07, P21) or romantic partners (P09).

Echoing their ESM responses, the interviewees also reported a high similarity between the final two response stages. First, as compared to the first two stages, being engaged with and acting upon notifications required that more physical and cognitive effort be expended on interacting with their devices. Consequently, they were unwilling or unable to commit to the remainder of the response process when it would be inconvenient for them to sustain that commitment. For instance, P01, P14, P22 and P25 told us that they delayed the Engage stage because they wanted more time to think about how to craft the message, or to actually craft it; and that they therefore delayed proceeding to the final two stages until they found themselves in situations that allowed them to do so, such as commuting by public transportation (P07). One commonly mentioned reason the interviewees gave for waiting for such moments to proceed the two latter response stages together, instead of separately, was to prevent them from forgetting to act upon notifications after reading them. As P07 said, *"I can't deal with the notification right away, but I still want to handle it later. So I just leave it there [in the notification bar] in case I forget to handle it"*. Leaving a notification unread provided these interviewees with a marker or cue to remind them to tap into it later.

In particular, IM messages' "seen" feature appeared to bind the final two stages of the response process together. The interviewees reported avoiding entering the message page specifically so that the "seen" sign, indicating that a message had been read, would not be shown to its sender. And, because tapping an IM notification will directly lead users to the message page, they avoid tapping such notifications as well. As P26 explained: *"I don't like to be left on 'seen' myself, so I avoid doing it to others. [...]. I usually respond to the message if I tap it."* An underlying reason for these positions is the potential social meaning of leaving someone a "seen" receipt, as articulated by P08: *"Since leaving someone on 'seen' is a way of stating your position nowadays, the recipient might take the message wrong."* Unlike the Glance stage, in which attention-switching is transient, Engage's requirement to commit one's entire attention to the phone during a social interaction was regarded as a social breach (P09, P11, P12, P18), especially when inattention to that interaction would be easily noticed. As P07 pointed out: *"When there were only a few people around, others can easily tell when I'm not paying attention to them."* In these situations, the interviewees tended not to initiate this stage of response, but thought it acceptable to glance at notifications (P02, P07), perhaps because it did not entail much physical movement (P10). Several participants also mentioned that privacy played an important role: i.e., they would avoid further interaction with their phones when in the midst of people to avoid their screen being seen by them (P06, P13, P28).

6 DISCUSSION

This paper's core findings are structured around two main themes. Firstly, by investigating users' perceived opportune moments for the four response stages, we elucidate the similarities and differences in how users perceive these moments across each stage. Secondly, through the exploration of the influence of various contextual factors on the perceived opportuneness of proceeding through each stage, we not only show various kinds of influence of contextual factors, but also reveal certain factors' influences on only specific response stages. A detailed discussion of each strand is provided in turn below.

6.1 The Opportune Moments for Engage and Act are Similar, whereas for Alert and Glance are Distinct

Drawing on both quantitative and qualitative findings, we observed distinct perceptions among participants regarding the opportune moments for the initial stages—Alert and Glance—compared to the subsequent stages. Conversely, the opportune moments for the latter stages—Engage and Act—exhibited notable similarities. This differentiation underscores the complexity of user interactions with notifications and suggests the need for stage-specific considerations in system design and user engagement strategies.

Specifically, our analysis revealed that the perceived opportune moments for all stages were correlated with each other and with overall perceptions of opportune moments. This suggests that these moments often co-occurred from the participants' perspective, indicating a general perception of opportune or inopportune times across all stages.

Despite this, our results showed varying degrees of correlation among the stages. Participants most commonly associated the general conception of an opportune moment for notifications with the moments for the Engage and Act stages. Moreover, the finding that the influences of contextual factors on the Engage and Act stages were more similar to each other than to any other pair of stages further supports the similarity between these two stages. Thus, while prior research has explored the concept of the opportune moment for different purposes, such as reducing disturbance [89] or enhancing receptivity to notifications [50], our results suggest that a "general concept" of an opportune moment for notifications, without specifying the response stage, may be more closely associated with engagement and acting than with generating an alert or glancing at notifications, at least from our participants' perspective. Nonetheless, given the sometimes simultaneous occurrence of these stages, as indicated by the correlations, the boundaries between them may not always be clear.

What we found particularly intriguing for future research is the potential impact on the proximity between the Engage and Act stages in a notification system designed to alleviate users' concerns about forgetting to act after engaging with content. This concern was especially prominent among participants dealing with messaging notifications, where forgetting to respond after reading could have social consequences, as also reported by Chou et al. [20]. If systems could offer cues, reminders, or easy access to act on notifications after reading them, this might reduce users' worries about forgetting. Then, it would be interesting to examine whether such system enhancements would effectively alleviate user concerns and encourage decoupling of the Engage and Act stages.

In contrast to the latter stages, the Alert and Glance stages were perceived as more distinct. Specifically, the Alert stage was notable in that participants' perceived opportune moments for this stage showed the least correlation with those for the other stages. Additionally, it was influenced by a unique set of contextual factors. Echoing results from prior work [14, 28, 89], which showed that controlling the ringer mode of the phone is sometimes intended to prevent disruptions in the environment, our participants also tended to associate the opportune moment for this stage with social-environmental factors, including the number of people interacting and social norms. The concern regarding the impact on the external entity—the surrounding, in addition to the notification recipient themselves—illustrates its distinctness in opportune moments from the other stages.

In contrast, the distinctiveness of the Glance stage is not primarily attributed to a specific set of contextual factors influencing it, but rather to the relatively low concern users have about the negative impacts of various contextual factors on proceeding to this stage. This is evidenced by the Glance stage being the only stage to receive an overall positive influence score on the opportuneness of the moment for proceeding, compared to all other stages. Previous research has documented users' high tendency to check notifications (e.g., [14, 23]) and the various reasons they browse their notifications [11]. Our study provides an explanation that bridges these observations. Specifically, it suggests that the frequent and prevalent occurrence of notification glances may not only stem from the various reasons people have for browsing notifications but also from the relatively lower threshold for a moment to be considered opportune for proceeding to this stage, combined with the perception that this stage is less subject to interference from various contextual factors compared to other stages.

Taken together, our results not only resonate with the division of interaction stages for notifications as proposed in prior research [29, 91], but also provide deeper insights into how the four stages were distinct from the lens of opportune moments. In light of these findings, we recommend that future notification systems predict opportune moments for these stages distinctly. Our results elucidate factors that are perceived both positively and negatively affecting the progression to these stages, as well as how these stages distinct from each other. Hence, we recommend future work consider these insights when designing opportune-moment detection mechanisms. For instance, notifications that require alerting users to draw their attention should consider environmental factors. Conversely, for notifications that necessitate actions, the system should avoid sending it when the user is busy or performing a complex task.

6.2 The Influence of Contextual Factors on Opportune Moments for the Four Stage is a Complex Phenomenon

Our second research question explores how the opportuneness of moments for each of the four response stages is impacted by various contextual factors. According to our results, we realize that addressing this question proves challenging due to the high complexity of the phenomenon, which manifests at least in four distinct elements identified in our study: 1) the status of the contextual factor (i.e., positive vs. negative, or high vs. low); 2) the specific response stage (i.e., Alert, Glance, Engage, and Act); 3) the type of notification (naively classified in this study as sender vs. non-sender); and 4) the types of influence measured (influence magnitude vs. influence likelihood). Through an analysis characterized by these elements, we elucidate general trends in influence scores and likelihoods, revealing how contextual factors can positively or negatively impact opportune moments for different stages. However, fully comprehending this phenomenon requires a thorough comparison among factors and stages, best visualized through the comprehensive diagram provided in the Appendix. This visualization aids in approaching the core of the phenomenon, but due to space limitations, this paper can only offer selective instances demonstrating the specific relationships between contextual factors and stages. For instance, our results include examples where certain contextual factors exhibit a type of influence on a specific stage only when in a certain status (e.g. activity-related factors, social-environmental factors). Furthermore, we observe variations in the prevalence and magnitude of influences: while some factors, like Happiness, have relatively low prevalence but average high influence scores, others, such as Content Attraction, show the opposite trend. The type of notification also plays a crucial role; content factors tend to have more negative impacts on the opportuneness of moments for notifications without a sender, yet exhibit positive influences for notifications with a sender.

All in all, building upon prior research on opportune moments (e.g. [15, 61, 66–68, 71, 74, 79, 88, 92]), breakpoints (e.g., [30, 38, 38, 69]), interpretability (e.g., [50, 63, 72, 105]), receptivity (e.g., [31, 61, 91]), and notification management (e.g., [60]), our study provides an important step towards understanding how opportune moments for different notification interactions are related to the studied contextual factors, and how and why such factors' influences can vary among response stages. Yet, the results also open the door to more questions requiring future investigation: such as "What constitutes an opportune moment?"; "To what extent can the studied contextual factors predict the opportuneness of moments for each response stage?"; "How well do these findings apply to opportune moments for other actions and more specific content, such as reading suggested materials [24, 74], interacting with a smart agent [58], responding to an intervention prompt [49, 53, 82], answering a questionnaire [33, 74], or performing a crowdsourcing task [13, 17, 18]?"

7 Research Limitation

This paper is subject to several limitations. First, its study design was inherently incapable of capturing people's usage of desktop/web applications or wearable devices, which could influence their actual interaction with notifications. For instance, a participant might choose not to respond to an IM notification on their smartphone if they already addressed the relevant message on their laptop. However, our ESM questionnaires did not ask participants to report information about their activity on other devices. Therefore, we could not gauge the extent to which such devices could have influenced their perceptions of opportune moments for each stage. Second, we only asked our participants how much each factor influenced opportune moments for each response stage, and not about opportune moments for notifications generally. In consequence, we were not able to fully explain the impact of each contextual factor on the general opportuneness. Third, given that participants are more likely to respond to ESM questionnaires when they are interruptible, and considering the duration of the 14-day study, it is likely that participants may feel fatigued and less inclined to answer the ESM questionnaires in the latter part of the study. Therefore, our dataset is very likely to be biased. Fourth, because we focused on response behavior in general, we did not study the interplay of our chosen factors with one another. For instance, it is intuitive that busyness would affect how smartphone users interact with notifications that take time to act upon. Future studies could, therefore, usefully investigate how various combinations of factors influence users. Fifth, the research app displayed the content of notifications in the Experience Sampling Method (ESM questionnaire. This likely made participants reluctant to respond to certain ESM questionnaires due to privacy concerns—specifically, the unwillingness to share notifications containing private, sensitive, or confidential information with the research team. This could potentially skew the results. It is likely that participants might perceive the appropriateness of the moment for handling such notifications differently from other notifications, but our dataset may not have captured these instances in the study. Sixth, as this study is exploratory in nature, we have employed Bonferroni correction to adjust our statistical significance level. Nevertheless, it's essential to recognize that while this correction reduces the risk of Type I errors, it simultaneously increases the chance of Type II errors (false negatives), potentially leading to genuinely significant differences being considered as non-significant [73]. Seventh, due to the high complexity of our data, the granularity of our findings is limited. This is, we used a relatively simple binary categorization of notifications into S- and NS-notifications, rather than a multi-class categorization like those adopted by app stores: into social media, games, messaging, shopping, and so on. We also classified the valence of a factor in only two ways, positive and negative, rather than, for example, a five-class valence system (high-positive, low-positive, neutral, low-negative, high-negative). However, doing so would have reduced the number of samples in each category and complicated our visualization/presentation of the influences of all factors. Nevertheless, with a sufficiently large data set, it might allow for more nuanced comparisons among these categories. Eighth, we conducted our field study in Taiwan; most of our participants were in their twenties; and half of them were students. As a result, it is unclear whether our findings can be generalized to another group of people from other age groups and cultures, or who have different IM-application preferences.

8 Conclusion

In this paper, we expanded the concept of the opportune moment by adopting a multi-step model that accounts for four stages of the notification-response process. Our ESM study with 74 participants yielded three high-level takeaways. First, we provided qualitative and quantitative evidence showing how the four stages are similar and distinct. For example, we demonstrated that the opportune moments for the Alert and Glance stages are relatively distinct from each other, whereas the moments for the final two stages, Engage and Act, are more closely aligned. Second, our results offer a more comprehensive understanding of how the perceived opportuneness of moments for each stage is related to various contextual factors. Specifically, we illustrated that these factors play varied roles in influencing which moments are perceived as opportune for receiving and addressing notifications, with their influences varying according to their statuses, whether the notification involves a sender, and the stage to which recipients need to proceed. However, given the highly complex interrelationships between factors, response stages, and notifications, our research findings represent an important yet initial step towards understanding this complex phenomenon. Building upon this work, we encourage further research to conduct cross-factor and cross-stage comparisons and analyses that extend beyond the scope of this paper.

9 Acknowledgement

We gratefully acknowledge the support provided by the National Science and Technology Council, Taiwan (110-2222-E-A49-008-MY3). We also deeply thank all the study participants, anonymous reviewers, and our colleagues, whose support and feedback contributed to the completion of this research.

References

[1] Piotr D. Adamczyk and Brian P. Bailey. 2004. If not now, when?: the effects of interruption at different moments within task execution. In *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*. ACM Press, Vienna, Austria, 271–278. https://doi.org/10.1145/985692.985727

[2] Richard A Armstrong. 2014. When to use the B onferroni correction. *Ophthalmic and Physiological Optics* 34, 5 (2014), 502–508.

[3] Daniel Avrahami and Scott E. Hudson. 2006. Communication characteristics of instant messaging: effects and predictions of interpersonal relationships. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work - CSCW '06*. ACM Press, Banff, Alberta, Canada, 505. https://doi.org/10.1145/1180875.1180954

[4] Daniel Avrahami and Scott E. Hudson. 2006. Responsiveness in instant messaging: predictive models supporting inter-personal communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. Association for Computing Machinery, New York, NY, USA, 731–740. https://doi.org/10.1145/1124772.1124881

[5] Brian Bailey and Joseph Konstan. 2006. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior* 22 (July 2006), 685–708. https://doi.org/10.1016/j.chb.2005.12.009

[6] B.P. Bailey, J.A. Konstan, and J.V. Carlis. 2000. Measuring the effects of interruptions on task performance in the user interface. In *Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0*, Vol. 2. 757–762 vol.2. https://doi.org/10.1109/ICSMC.2000.885940 ISSN: 1062-922X.

[7] Nikola Banovic, Christina Brant, Jennifer Mankoff, and Anind Dey. 2014. ProactiveTasks: the short of mobile device use sessions. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services (MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 243–252. https://doi.org/10.1145/2628363.2628380

[8] Jiang Bian, A. Dong, Xiaofeng He, S. Reddy, and Yi Chang. 2013. User Action Interpretation for Online Content Optimization. *IEEE Transactions on Knowledge and Data Engineering* (2013). https://doi.org/10.1109/TKDE.2012.130

[9] Seyma Kucukozer Cavdar, Tugba Taskaya-Temizel, Mirco Musolesi, and Peter Tino. 2020. A Multi-perspective Analysis of Social Context and Personal Factors in Office Settings for the Design of an Effective Mobile Notification System. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (March 2020), 15:1–15:38. https://doi.org/10.1145/3381000

[10] Narae Cha, Auk Kim, Cheul Young Park, Soowon Kang, Mingyu Park, Jae-Gil Lee, Sangsu Lee, and Uichin Lee. 2020. Hello There! Is Now a Good Time to Talk? Opportune Moments for Proactive Interactions with Smart Speakers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (Sept. 2020), 74:1–74:28. https://doi.org/10.1145/3411810

[11] Xi-Jing Chang, Fang-Hsin Hsu, En-Chi Liang, Zih-Yun Chiou, Ho-Hsuan Chuang, Fang-Ching Tseng, Yu-Hsin Lin, and Yung-Ju Chang. 2023. Not Merely Deemed as Distraction: Investigating Smartphone Users' Motivations for Notification-Interaction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[12] Yung-Ju Chang, Yi-Ju Chung, and Yi-Hao Shih. 2019. I Think It's Her: Investigating Smartphone Users' Speculation about Phone Notifications and Its Influence on Attendance. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Taipei Taiwan, 1–13. https://doi.org/10.1145/3338286.3340125

[13] Yung-Ju Chang, Gaurav Paruthi, Hsin-Ying Wu, Hsin-Yu Lin, and Mark W Newman. 2017. An investigation of using mobile and situated crowdsourcing to collect annotated travel activity data in real-word settings. *International Journal of Human-Computer Studies* 102 (2017), 81–102.

[14] Yung-Ju Chang and John C. Tang. 2015. Investigating Mobile Users' Ringer Mode Usage and Attentiveness and Responsiveness to Communication. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Copenhagen Denmark, 6–15. https://doi.org/10.1145/2785830.2785852

[15] Kuan-Wen Chen, Yung-Ju Chang, and Liwei Chan. 2022. Predicting Opportune Moments to Deliver Notifications in Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 186, 18 pages. https://doi.org/10.1145/3491102.3517529

[16] Yu-Chun Chen, Yu-Jen Lee, Kuei-Chun Kao, Jie Tsai, En-Chi Liang, Wei-Chen Chiu, Faye Shih, and Yung-Ju Chang. 2023. Are you killing time? predicting smartphone users' time-killing moments via fusion of smartphone sensor data and screenshots. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[17] Justin Cheng, Jaime Teevan, Shamsi T. Iqbal, and Michael S. Bernstein. 2015. Break It Down: A Comparison of Macro- and Microtasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 4061–4064. https://doi.org/10.1145/2702123.2702146

[18] Chia-En Chiang, Yu-Chun Chen, Fang-Yu Lin, Felicia Feng, Hao-An Wu, Hao-Ping Lee, Chang-Hsuan Yang, and Yung-Ju Chang. 2021. "I Got Some Free Time": Investigating Task-execution and Task-effort Metrics in Mobile Crowdsourcing Tasks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445477

[19] Yu-Ling Chou, Yu-Ling Chien, Yu-Hsin Lin, Kung-Pai Lin, Faye Shih, and Yung-Ju Chang. 2022. Because I'm Restricted, 2–4 PM Unable to See Messages: Exploring Users' Perceptions and Likely Practices around Exposing Attention Management Use on IM Online Status. In *CHI Conference on Human Factors in Computing Systems*. 1–18.

[20] Yu-Ling Chou, Yi-Hsiu Lin, Tzu-Yi Lin, Hsin Ying You, and Yung-Ju Chang. 2022. Why Did You/I Read but Not Reply? IM Users' Unresponded-to Read-receipt Practices and Explanations of Them. In *CHI Conference on Human Factors in Computing Systems*. 1–15.

[21] Ed Diener and Louis Tay. 2014. Review of the Day Reconstruction Method (DRM). *Social Indicators Research* 116, 1 (March 2014), 255–267. https://doi.org/10.1007/s11205-013-0279-x

[22] Tilman Dingler and Martin Pielot. 2015. I'll be there for you: Quantifying Attentiveness towards Mobile Messaging. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–5.

[23] Tilman Dingler and Martin Pielot. 2015. I'll be there for you: Quantifying Attentiveness towards Mobile Messaging. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 1–5. https://doi.org/10.1145/2785830.2785840

[24] Tilman Dingler, Benjamin Tag, Sabrina Lehrer, and Albrecht Schmidt. 2018. Reading Scheduler: Proactive Recommendations to Help Users Cope with Their Daily Reading Volume. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM 2018)*. Association for Computing Machinery, New York, NY, USA, 239–244. https://doi.org/10.1145/3282894.3282917

[25] Samantha Dockray, Nina Grant, Arthur A. Stone, Daniel Kahneman, Jane Wardle, and Andrew Steptoe. 2010. A Comparison of Affect Ratings Obtained with Ecological Momentary Assessment and the Day Reconstruction Method. *Social Indicators Research* 99, 2 (Nov. 2010), 269–283. https://doi.org/10.1007/s11205-010-9578-7

[26] Fiona Draxler, Julia Maria Brenner, Manuela Eska, Albrecht Schmidt, and Lewis L Chuang. 2022. Agenda- and Activity-Based Triggers for Microlearning. In *27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 620–632. https://doi.org/10.1145/3490099.3511133

[27] Anja Exler, Marcel Braith, Kristina Mincheva, Andrea Schankin, and Michael Beigl. 2018. Smartphone-Based Estimation of a User Being in Company or Alone Based on Place, Time, and Activity. In *Mobile Computing, Applications, and Services*, Kazuya Murao, Ren Ohmura, Sozo Inoue, and Yusuke Gotoh (Eds.). Vol. 240. Springer International Publishing, Cham, 74–89. https://doi.org/10.1007/978-3-319-90740-6_5 Series Title: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering.

[28] Anja Exler, Zeynep Günes, and Michael Beigl. 2019. Preferred notification modalities depending on the location and the location-based activity. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. ACM, London United Kingdom, 1064–1069. https://doi.org/10.1145/3341162.3344842

[29] Tseng Fang-Ching, Chiou Zih-Yun, Chuang Ho-Hsuan, Su Li-Ting, Lin Yong-Han, Lin Yu-Rou, Lee Yi-Chi, Wang Peng-Jui, Chen Uei-Dar, and Chang Yung-Ju. 2023. Multiple Device Users' Actual and Ideal Cross-Device Usage for Multi-Stage Notifcation-Interactions: An ESM Study Addressing the Usage Gap and Impacts of Device Context. (2023).

[30] Joel E. Fischer, Chris Greenhalgh, and Steve Benford. 2011. Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '11*. ACM Press, Stockholm, Sweden, 181. https://doi.org/10.1145/2037373.2037402

[31] Joel E. Fischer, Nick Yee, Victoria Bellotti, Nathan Good, Steve Benford, and Chris Greenhalgh. 2010. Effects of content and time of delivery on receptivity to mobile interruptions. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*. ACM Press, Lisbon, Portugal, 103. https://doi.org/10.1145/1851600.1851620

[32] Claudio Forlivesi, Utku Günay Acer, Marc van den Broeck, and Fahim Kawsar. 2018. Mindful interruptions: a lightweight system for managing interruptibility on wearables. In *Proceedings of the 4th ACM Workshop on Wearable Systems and Applications (WearSys '18)*. Association for Computing Machinery, New York, NY, USA, 27–32. https://doi.org/10.1145/3211960.3211974

[33] Bo-Jhang Ho, Bharathan Balaji, Mehmet Koseoglu, Sandeep Sandha, Siyou Pei, and Mani Srivastava. 2020. Quick Question: Interrupting Users for Microtasks with Reinforcement Learning. *arXiv:2007.09515 [cs]* (July 2020). http://arxiv.org/abs/2007.09515 arXiv: 2007.09515.

[34] Joyce Ho and Stephen S. Intille. 2005. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*. ACM Press, Portland, Oregon, USA, 909. https://doi.org/10.1145/1054972.1055100

[35] Ching-Yu Hsieh, Yi-Shyuan Chiang, Hung-Yu Chiu, and Yung-Ju Chang. 2020. Bridging the Virtual and Real Worlds: A Preliminary Study of Messaging Notifications in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376228

[36] Scott Hudson, James Fogarty, Christopher Atkeson, Daniel Avrahami, Jodi Forlizzi, Sara Kiesler, Johnny Lee, and Jie Yang. 2003. Predicting human interruptibility with sensors: a Wizard of Oz feasibility study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. Association for Computing Machinery, New York, NY, USA, 257–264. https://doi.org/10.1145/642611.642657

[37] Shamsi Iqbal and Brian Bailey. 2006. Leveraging Characteristics of Task Structure to Predict the Cost of Interruption. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 2 (April 2006). https://doi.org/10.1145/1124772.1124882

[38] Shamsi T. Iqbal and Brian P. Bailey. 2008. Effects of intelligent notification management on users and their tasks. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, Florence, Italy, 93. https://doi.org/10.1145/1357054.1357070

[39] Shamsi T. Iqbal and Brian P. Bailey. 2011. Oasis: A framework for linking notification delivery to the perceptual structure of goal-directed tasks. *ACM Transactions on Computer-Human Interaction* 17, 4 (Dec. 2011), 15:1–15:28. https://doi.org/10.1145/1879831.1879833

[40] E. Isaacs, N. Yee, D. Schiano, Nathaniel Good, Nicolas Ducheneaut, and V. Bellotti. 2010. - 1-Mobile Microwaiting Moments : The Role of Context in Receptivity to Content While on the Go. https://www.semanticscholar.org/paper/1-Mobile-Microwaiting-Moments-%3A-The-Role-of-Context-Isaacs-Yee/cafaaa823351104bddb3f95071241866b3993dbe

[41] Daniel Kahneman, Alan B. Krueger, David A. Schkade, Norbert Schwarz, and Arthur A. Stone. 2004. A survey method for characterizing daily life experience: the day reconstruction method. *Science (New York, N.Y.)* 306, 5702 (Dec. 2004), 1776–1780. https://doi.org/10.1126/science.1103572

[42] Elizabeth A. Kensinger. 2009. Remembering the Details: Effects of Emotion. *Emotion review* 1, 2 (2009), 99–113. https://doi.org/10.1177/1754073908100432

[43] N. Kern, S. Antifakos, B. Schiele, and A. Schwaninger. 2004. A Model for Human Interruptability: Experimental Evaluation and Automatic Estimation from Wearable Sensors. In *Eighth International Symposium on Wearable Computers*. IEEE, Arlington, VA, USA, 158–165. https://doi.org/10.1109/ISWC.2004.3

[44] Muhammad Faizan Khan, Lu Lu, Muhammad Toseef, Ahmed Musyafa, and Ahmad Amin. 2022. NotifyMiner: rule based user behavioral machine learning approach for context wise personalized notification services. *Journal of Ambient Intelligence and Humanized Computing* (2022), 1–17.

[45] Jinhyuk Kim, Hiroe Kikuchi, and Yoshiharu Yamamoto. 2012. Systematic comparison between ecological momentary assessment and day reconstruction method for fatigue and mood states in healthy adults. *British journal of health psychology* 18 (Sept. 2012). https://doi.org/10.1111/bjhp.12000

[46] Andreas Komninos, Elton Frengkou, and John Garofalakis. 2018. Predicting user responsiveness to smartphone notifications for edge computing. In *European Conference on Ambient Intelligence*. Springer, 3–19.

[47] Kostadin Kushlev, Bruno Cardoso, and Martin Pielot. 2017. Too tense for candy crush: affect influences user engagement with proactively suggested content. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Vienna Austria, 1–6. https://doi.org/10.1145/3098279.3098569

[48] Kostadin Kushlev, Jason Proulx, and Elizabeth W. Dunn. 2016. "Silence Your Phones": Smartphone Notifications Increase Inattention and Hyperactivity Symptoms. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 1011–1020. https://doi.org/10.1145/2858036.2858359

[49] Neal Lathia, Veljko Pejovic, Kiran K. Rachuri, Cecilia Mascolo, Mirco Musolesi, and Peter J. Rentfrow. 2013. Smartphones for Large-Scale Behavior Change Interventions. *IEEE Pervasive Computing* 12, 3 (July 2013), 66–73. https://doi.org/10.1109/MPRV.2013.56 Conference Name: IEEE Pervasive Computing.

[50] Hao-Ping Lee, Kuan-Yin Chen, Chih-Heng Lin, Chia-Yu Chen, Yu-Lin Chung, Yung-Ju Chang, and Chien-Ru Sun. 2019. Does *Who* Matter? Studying the Impact of Relationship Characteristics on Receptivity to Mobile IM Messages. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300756

[51] Hao-Ping Lee, Tilman Dingler, Chih-Heng Lin, Kuan-Yin Chen, Yu-Lin Chung, Chia-Yu Chen, and Yung-Ju Chang. 2019. Predicting smartphone users' general responsiveness to IM contacts based on IM behavior. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–6.

[52] Hao-Ping Hank Lee, Yi-Shyuan Chiang, Yu-Ling Chou, Kung-Pai Lin, and Yung-Ju Chang. 2023. What makes IM users (un) responsive: An empirical investigation for understanding IM responsiveness. *International Journal of Human-Computer Studies* 172 (2023), 102983.

[53] Aniek Lentferink, Matthijs L. Noordzij, Anouk Burgler, Randy Klaassen, Youri Derks, Hilbrand Oldenhuis, Hugo Velthuijsen, and Lisette van Gemert-Pijnen. 2021. On the receptivity of employees to just-in-time self-tracking and eCoaching for stress management: a mixed-methods approach. *Behaviour & Information Technology* 0, 0 (Feb. 2021), 1–27. https://doi.org/10.1080/0144929X.2021.1876764 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/0144929X.2021.1876764.

[54] Tzu-Chieh Lin, Yu-Shao Su, Emily Helen Yang, Yun Han Chen, Hao-Ping Lee, and Yung-Ju Chang. 2021. “Put it on the Top, I’ll Read it Later”: Investigating Users’ Desired Display Order for Smartphone Notifications. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 520. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3411764.3445384

[55] Tzu-Chieh Lin, Yu-Shao Su, Emily Helen Yang, Yun Han Chen, Hao-Ping Lee, and Yung-Ju Chang. 2021. "Put it on the Top, I'll Read it Later": Investigating Users' Desired Display Order for Smartphone Notifications. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13. https://doi.org/10.1145/3411764.3445384

[56] Richard E. Lucas, Carol Wallsworth, Ivana Anusic, and M. Brent Donnellan. 2021. A direct comparison of the day reconstruction method (DRM) and the experience sampling method (ESM). *Journal of Personality and Social Psychology* 120, 3 (March 2021), 816–835. https://doi.org/10.1037/pspp0000289

[57] Gloria Mark, Shamsi Iqbal, Mary Czerwinski, and Paul Johns. 2015. Focused, Aroused, but so Distractible: Temporal Perspectives on Multitasking and Communications. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, Vancouver BC Canada, 903–916. https://doi.org/10.1145/2675133.2675221

[58] D. McFarlane. 1999. Coordinating the Interruption of People in Human-Computer Interaction. *undefined* (1999). https://www.semanticscholar.org/paper/Coordinating-the-Interruption-of-People-in-McFarlane/a4b595558f76fc5449d762a49e6a651af121b9d1

[59] Daniel C McFarlane. 1999. Coordinating the interruption of people in human-computer interaction.. In *Interact*, Vol. 99. 295.

[60] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. 2016. PrefMiner: mining user's preferences for intelligent mobile notification management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 1223–1234. https:

//doi.org/10.1145/2971648.2971747

[61] Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. 2015. Designing content-driven intelligent notification mechanisms for mobile applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, Osaka, Japan, 813–824. https://doi.org/10.1145/2750858.2807544

[62] Abhinav Mehrotra, Sandrine R. Müller, Gabriella M. Harari, Samuel D. Gosling, Cecilia Mascolo, Mirco Musolesi, and Peter J. Rentfrow. 2017. Understanding the Role of Places and Activities on Mobile Phone Interaction and Usage Patterns. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (Sept. 2017), 1–22. https://doi.org/10.1145/3131901

[63] Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. 2016. My Phone and Me: Understanding People's Receptivity to Mobile Notifications. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 1021–1032. https://doi.org/10.1145/2858036.2858566

[64] Abhinav Mehrotra, Fani Tsapeli, Robert Hendley, and Mirco Musolesi. 2017. MyTraces: Investigating Correlation and Causation between Users’ Emotional States and Mobile Phone Interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (Sept. 2017), 83:1–83:21. https://doi.org/10.1145/3130948

[65] Tamir Mendel, Roei Schuster, Eran Tromer, and Eran Toch. 2022. Toward Proactive Support for Older Adults: Predicting the Right Moment for Providing Mobile Safety Help. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1, Article 25 (mar 2022), 25 pages. https://doi.org/10.1145/3517249

[66] Tadashi Okoshi, Julian Ramos, Hiroki Nozaki, Jin Nakazawa, Anind K. Dey, and Hideyuki Tokuda. 2015. Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, St. Louis, MO, USA, 96–104. https://doi.org/10.1109/PERCOM.2015.7146515

[67] Tadashi Okoshi, Kota Tsubouchi, and Hideyuki Tokuda. 2018. Real-world large-scale study on adaptive notification scheduling on smartphones. *Pervasive and Mobile Computing* 50 (Oct. 2018), 1–24. https://doi.org/10.1016/j.pmcj.2018.07.005

[68] Tadashi Okoshi, Kota Tsubouchi, and Hideyuki Tokuda. 2019. Real-World Product Deployment of Adaptive Push Notification Scheduling on Smartphones. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*. Association for Computing Machinery, New York, NY, USA, 2792–2800. https://doi.org/10.1145/3292500.3330732

[69] Chunjong Park, Junsung Lim, Juho Kim, Sung-Ju Lee, and Dongman Lee. 2017. Don't Bother Me. I'm Socializing! A Breakpoint-Based Smartphone Notification System. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 541–554. https://doi.org/10.1145/2998181.2998189

[70] Sameer Patil, Roberto Hoyle, Roman Schlegel, Apu Kapadia, and Adam J. Lee. 2015. Interrupt Now or Inform Later? Comparing Immediate and Delayed Privacy Feedback. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1415–1418. https://doi.org/10.1145/2702123.2702165

[71] Veljko Pejovic and Mirco Musolesi. 2014. InterruptMe: designing intelligent prompting mechanisms for pervasive applications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. Association for Computing Machinery, New York, NY, USA, 897–908. https://doi.org/10.1145/2632048.2632062

[72] Veljko Pejovic, Mirco Musolesi, and Abhinav Mehrotra. 2015. Investigating The Role of Task Engagement in Mobile Interruptibility. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, Copenhagen Denmark, 1100–1105. https://doi.org/10.1145/2786567.2794336

[73] Thomas V Perneger. 1998. What's wrong with Bonferroni adjustments. *Bmj* 316, 7139 (1998), 1236–1238.

[74] Martin Pielot, Bruno Cardoso, Kleomenis Katevas, Joan Serrà, Aleksandar Matic, and Nuria Oliver. 2017. Beyond Interruptibility: Predicting Opportune Moments to Engage Mobile Phone Users. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (Sept. 2017), 1–25. https://doi.org/10.1145/3130956

[75] Martin Pielot, Karen Church, and Rodrigo de Oliveira. 2014. An in-situ study of mobile phone notifications. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services - MobileHCI '14*. ACM Press, Toronto, ON, Canada, 233–242. https://doi.org/10.1145/2628363.2628364

[76] Martin Pielot, Rodrigo De Oliveira, Haewoon Kwak, and Nuria Oliver. 2014. Didn't you see my message? predicting attentiveness to mobile instant messages. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3319–3328.

[77] Martin Pielot, Tilman Dingler, Jose San Pedro, and Nuria Oliver. 2015. When attention is not scarce - detecting boredom from mobile phone usage. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*. ACM Press, Osaka, Japan, 825–836. https://doi.org/10.1145/2750858.2804252

[78] Martin Pielot, Amalia Vradi, and Souneil Park. 2018. Dismissed!: a detailed exploration of how mobile phone users handle push notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Barcelona Spain, 1–11. https://doi.org/10.1145/3229434.3229445

[79] Benjamin Poppinga, Wilko Heuten, and Susanne Boll. 2014. Sensor-Based Identification of Opportune Moments for Triggering Notifications. *IEEE Pervasive Computing* 13, 1 (2014), 22–29. https://doi.org/10.1109/MPRV.2014.15

[80] Stephanie Rosenthal, Anind K. Dey, and Manuela Veloso. 2011. Using Decision-Theoretic Experience Sampling to Build Personalized Mobile Phone Interruption Models. In *Pervasive Computing*, Kent Lyons, Jeffrey Hightower, and Elaine M. Huang (Eds.). Vol. 6696. Springer Berlin Heidelberg, Berlin, Heidelberg, 170–187. https://doi.org/10.1007/978-3-642-21726-5_11 Series Title: Lecture Notes in Computer Science.

[81] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. 2014. Large-scale assessment of mobile notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 3055–3064. https://doi.org/10.1145/2556288.2557189

[82] Akane Sano, Paul Johns, and Mary Czerwinski. 2017. Designing opportune stress intervention delivery timing using multi-modal data. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. 346–353. https://doi.org/10.1109/ACII.2017.8273623 ISSN: 2156-8111.

[83] Hillol Sarker, Moushumi Sharmin, Amin Ahsan Ali, Md. Mahbubur Rahman, Rummana Bari, Syed Monowar Hossain, and Santosh Kumar. 2014. Assessing the availability of users to engage in just-in-time intervention in the natural environment. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, Seattle Washington, 909–920. https://doi.org/10.1145/2632048.2636082

[84] Holger Schielzeth, Niels J Dingemanse, Shinichi Nakagawa, David F Westneat, Hassen Allegue, Céline Teplitsky, Denis Réale, Ned A Dochtermann, László Zsolt Garamszegi, and Yimen G Araya-Ajoy. 2020. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in ecology and evolution* 11, 9 (2020), 1141–1152.

[85] Florian Schulze and Georg Groh. 2014. Studying how character of conversation affects personal receptivity to mobile notifications. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 1729–1734. https://doi.org/10.1145/2559206.2581320

[86] Florian Schulze and Georg Groh. 2016. Conversational context helps improve mobile notification management. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Florence Italy, 518–528. https://doi.org/10.1145/2935334.2935347

[87] Alireza Sahami Shirazi and Niels Henze. 2015. Assessment of Notifications on Smartwatches. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, Copenhagen Denmark, 1111–1116. https://doi.org/10.1145/2786567.2794338

[88] G. H. (Henri) ter Hofte. 2007. Xensible interruptions from your mobile phone. In *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services - MobileHCI '07*. ACM Press, Singapore, 178–181. https://doi.org/10.1145/1377999.1378003

[89] Eran Toch, Hadas Chassidim, and Tali Hatuka. 2020. Can you Turn it Off?: The Spatial and Social Context of Mobile Disturbance. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020), 1–18. https://doi.org/10.1145/3415162

[90] J. Gregory Trafton, Erik M Altmann, Derek P Brock, and Farilee E Mintz. 2003. Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies* 58, 5 (May 2003), 583–603. https://doi.org/10.1016/S1071-5819(03)00023-5

[91] Liam D. Turner, Stuart M. Allen, and Roger M. Whitaker. 2015. Push or Delay? Decomposing Smartphone Notification Response Behaviour. In *Proceedings of the 6th International Workshop on Human Behavior Understanding - Volume 9277*. Springer-Verlag, Berlin, Heidelberg, 69–83. https://doi.org/10.1007/978-3-319-24195-1_6

[92] Liam D. Turner, Stuart M. Allen, and Roger M. Whitaker. 2017. Reachable but not receptive: Enhancing smartphone interruptibility prediction by modelling the extent of user engagement with notifications. *Pervasive and Mobile Computing* 40 (Sept. 2017), 480–494. https://doi.org/10.1016/j.pmcj.2017.01.011

[93] Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The experience sampling method on mobile devices. *ACM Computing Surveys (CSUR)* 50, 6 (2017), 1–40.

[94] Niels van Berkel, Jorge Goncalves, Peter Koval, Simo Hosio, Tilman Dingler, Denzil Ferreira, and Vassilis Kostakos. 2019. Context-informed scheduling and analysis: improving accuracy of mobile self-reports. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[95] Niels van Berkel, Jorge Goncalves, Lauri Lovén, Denzil Ferreira, Simo Hosio, and Vassilis Kostakos. 2019. Effect of experience sampling schedules on response rate and recall accuracy of objective self-reports. *International Journal of Human-Computer Studies* 125 (2019), 118–128.

[96] Niels Van Berkel, Chu Luo, Theodoros Anagnostopoulos, Denzil Ferreira, Jorge Goncalves, Simo Hosio, and Vassilis Kostakos. 2016. A systematic assessment of smartphone usage gaps. In *Proceedings of the 2016 CHI conference on*

human factors in computing systems. 4711–4721.
[97] Aku Visuri, Niels van Berkel, Tadashi Okoshi, Jorge Goncalves, and Vassilis Kostakos. 2019. Understanding smartphone notifications' user interactions and content importance. *International Journal of Human-Computer Studies* 128 (Aug. 2019), 72–85. https://doi.org/10.1016/j.ijhcs.2019.03.001
[98] Peter Vorderer, Dorothée Hefner, Leonard Reinecke, and Christoph Klimmt (Eds.). 2018. *Permanently online, permanently connected: living and communicating in a POPC world.* Routledge, Taylor & Francis Group, New York ; London.
[99] Dominik Weber, Alexandra Voit, Gisela Kollotzek, and Niels Henze. 2019. Annotif: A System for Annotating Mobile Notifcations in User Studies. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia (MUM '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3365610.3365611
[100] Tilo Westermann, Ina Wechsung, and Sebastian Möller. 2016. Smartphone Notifications in Context: a Case Study on Receptivity by the Example of an Advertising Service. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 2355–2361. https://doi.org/10.1145/2851581.2892383
[101] Ting-Wei Wu, Yu-Ling Chien, Hao-Ping Lee, and Yung-Ju Chang. 2021. IM Receptivity and Presentation-type Preferences among Users of a Mobile App with Automated Receptivity-status Adjustment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. https://doi.org/10.1145/3411764.3445209
[102] Chang XiJing, Hsu Fang-Hsin, Liang En-Chi, Chiou Zih-Yun, Chuang Ho-Hsuan, Tseng Fang-Ching, Lin Yu-Hsin, and Chang Yung-Ju. 2023. Not Merely Deemed as Distraction: Investigating Smartphone Users' Motivations for Notification-Interaction. (2023).
[103] Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating How Experienced UX Designers Effectively Work with Machine Learning. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 585–596. https://doi.org/10.1145/3196709.3196730
[104] SungHyuk Yoon, Sang-su Lee, Jae-myung Lee, and KunPyo Lee. 2014. Understanding notification stress of smartphone messenger app. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 1735–1740. https://doi.org/10.1145/2559206.2581167
[105] Fengpeng Yuan, Xianyi Gao, and Janne Lindqvist. 2017. How Busy Are You?: Predicting the Interruptibility Intensity of Mobile Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5346–5360. https://doi.org/10.1145/3025453.3025946

Received February 2024; revised May 2024; accepted June 2024