Design guidelines for studies incorporating passive mobile sensing in the wild

ANONYMOUS AUTHOR(S)

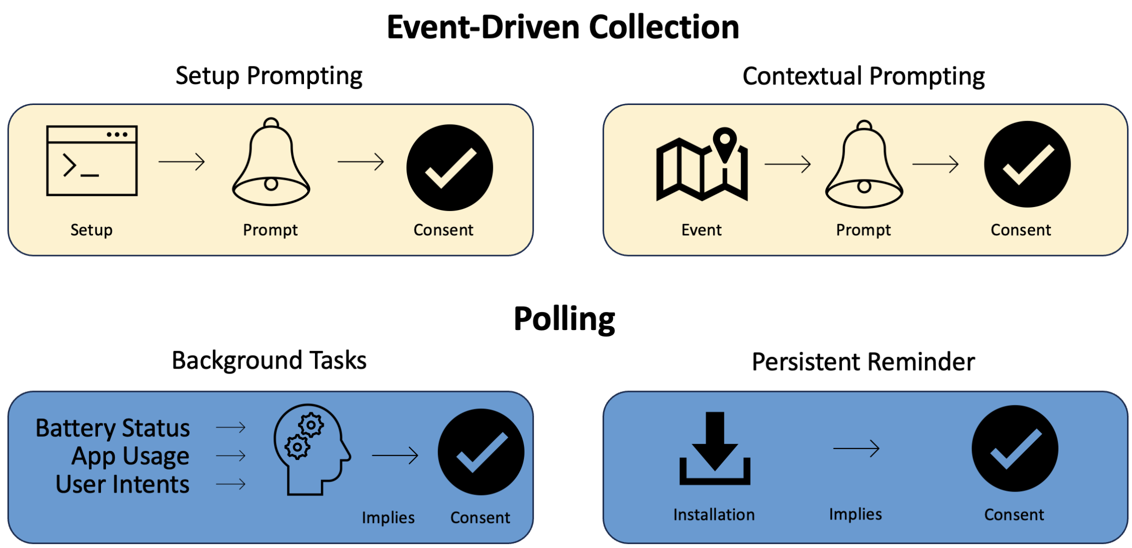

Fig. 1. We studied two methods of passive sensing: Event-driven collection and polling. For Event-driven collection we compared setup and contextual prompting authorization. For polling consent is implied. We compared traditional background tasks to a new approach using a persistent reminder. All four of these paradigms involve different models of authorization for passive mobile sensing.

Mobile devices are powerful tools for passively collecting health data used to train classifiers which can detect adverse health events, enabling real-time interventions. We study the authorization procedures for two paradigms for passive data collection: event-driven collection and polling. Event-driven authorization is facilitated through setup prompting or contextual prompting, whereas, polling via background tasks is authorized through an implied consent model. We also studied a new method: polling via persistent reminders. We built and tested a passive sensing framework through a mixed-method user study analyzing the effectiveness of these paradigms in the wild. For event-driven data collection, we found that contextual prompting was 55.46% more effective than setup prompting. The persistent reminder was 795% more effective for polling than background tasks on iOS devices. Using these findings and qualitative insights derived from a survey of participants, we formalize a set of design guidelines for passive mobile sensing in the wild.

CCS Concepts: • **Human-centered computing** → **Mobile computing**; **User studies**; • **Security and privacy** → **Usability in security and privacy**; • **Applied computing** → **Consumer health**.

Additional Key Words and Phrases: mobile health sensing, ecological momentary assessment, experience sampling, passive data collection, mobile permissions, notification systems

ACM Reference Format:
Anonymous Author(s). 2024. Design guidelines for studies incorporating passive mobile sensing in the wild. In *Proceedings of CHI 24, June 03–05, 2024, Honolulu, HI*. ACM, New York, NY, USA, 15 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Research at the intersection of health and Human-Computer Interaction (HCI) frequently uses mobile and ubiquitous devices as valuable tools for gathering health data in real-world settings [11, 20, 21, 24, 32, 40, 43–45]. In its early stages, mobile devices were predominantly used for gathering ecological momentary assessments (EMA) [9, 18, 19], a methodology that entails soliciting user assessments within natural environments. With the evolution of mobile devices, incorporating an increasing number of sensors and expanding their capabilities to collect and access health-related data, researchers have also embraced passive data collection methods. Passive data collection involves the gathering of information without requiring active user involvement. This can be achieved through two main paradigms: polling and event-driven collection. Polling uses background runtime to gather sensor data and logs, while event-driven collection is triggered by an event, like location changes, capturing the event with other sensor data and logs.

In many research studies, a combination of ecological momentary assessments (EMA) and passive sensing techniques are employed. This integration enables researchers to leverage the benefits of both approaches, providing a more comprehensive understanding of users’ experiences and behaviors in real-world contexts [35]. Passively recorded sensor data streams are often used as inputs to machine learning models that predict outcomes captured through EMA [1, 3, 16, 27].

However, the objectives of mobile operating systems to protect user privacy and enhance battery performance are frequently at odds with the goals of passive data collection studies. Passive data collection carries the risk of compromising user privacy and requires background runtime, which can lead to reduced battery life. Most research in passive data collection has been conducted using the Android operating system due to its more permissive API and fewer restrictions on background runtime [7], but there has been some progress towards creating passive mobile data collection systems for iPhone (iOS) devices [29]. With devices continuing to add privacy and performance features, further efforts are needed to formalize design guidelines that align with the mobile OS’ objectives to safeguard users’ privacy and optimize battery performance while still enabling the flexibility required for longitudinal passive sensing studies [36].

To develop design guidelines for in-the-wild passive sensing, we developed a passive sensing framework that employs EMA, polling, and event-driven data collection. The framework consists of Android and iOS applications integrated with a cloud-based storage system. To cooperate with the mobile devices’ privacy and battery performance goals, our framework tested the various authorization procedures provided by the mobile OSs, shown in Figure 1.

For event-driven collection, users need to grant authorization to access background events. Both Android and iOS implement ask-on-first-use promots [47], where permission is requested when the user first uses or activates the feature that requires access to the event. With *setup prompting*, implemented on Android devices, authorization needs to be requested before the event occurs. For passive sensing apps, this usually occurs during setup. *Contextual prompting* on iOS devices prompts the user when the first event occurs.

Polling is usually accomplished with background tasks which are granted runtime through a machine-learning model. The ML model considers app usage, battery status, the presence of the app in the app switcher, and other factors to determine when background runtime should be granted. Although Android devices tend to be more permissive than iOS devices, consistency varies across specific Android models. Due to this inconsistency, our framework also introduced a persistent reminder, or home screen widget, as an alternative polling method. Installation of the widget implies consent for passive data collection.

In addition to passive data collection, many sensing apps also incorporate EMA which relies on notifications to inform participants to complete assessments. However, mobile devices have added features, such as notification summaries, distraction-free time, and authorization through contextual prompts to their notification systems.

Using our framework, we tested the feasibility of passively collecting data in a user study to answer the following questions about passive sensing in the wild:

RQ1: For event-driven data collection, how effective are the setup and contextual prompting paradigms in granting authorization to recurrent background events on both a systems and user level?

RQ2: For polling-based data collection, how effective are persistent reminders compared to traditional background tasks on both a systems and user level?

RQ3: How does the addition of supplemental user-interface decisions such as notification summaries, distraction-free times, and contextual prompting affect the success of EMA notification delivery and corresponding user perceptions?

Using the results of our user study, we contributed the following:

- We studied the effectiveness of different authorization processes for event-driven data collection methods.
- We studied the use of a persistent reminder as an alternative method to polling and compared it to background tasks.
- Since EMA relies on notifications to inform users to complete assessments, we also studied the effectiveness of setup and contextual prompting for authorization to send notifications.
- We developed design guidelines to help researchers optimize passive sensing frameworks' ability to collect health data.

2 RELATED WORK

2.1 Ecological Momentary Assessment

EMA, also know as experience sampling, is increasingly applied in fields of HCI research [26, 38], health studies and interventions [14, 22, 34], and crowdsourcing [10, 12, 17, 20]. Shiffman et al. showed that EMA can advance "the science and practice of clinical psychology by shedding light on the dynamics of behavior in real-world settings" [34]. Even though EMA has a long history there are still many technical, methodological, privacy, compliance, and psychometric challenges [13, 36, 39].

Notifications play a primary role in alerting and reminding users to complete EMA tasks. However, Yoon et al. [48] showed that notifications can cause stress in users and Kane et al. [22] show that EMA notifications could alter the emotional state of participants. To reduce the burden of completing EMA tasks, Chan et al. [8] introduced a micro-interaction notification system to collect EMAs. To improve notification receptibility, Li et al. [26] developed personalized models to understand and conform to notification preferences. Mobile devices are continuously enhancing their notification systems to assist users in controlling interruptions, including notification deferral, suppression, and snoozing [4]. In addition, iOS devices use contextual permission prompting to authorize notifications. We contribute to this work through qualitative analysis of EMA notifications delivered within current notification systems.

2.2 Passive Mobile Health Sensing

Passive mobile health sensing is collecting data through the use of smart mobile devices, such as smartphones, smart-watches, and other mobile sensors without the participant actively participating in data collection [2, 25, 42]. For example, Jongs et al. used smartphone-based location data to access neuropsychiatric phenotypes [21], Doryab et al. used a combination of call logs, location, screen time, and activity data to predict social isolation [16] and Rabbi et al. used the mobile sensing to predict mental health problems [31]. Norbi et al. used a combination of social media and smartphone app usage to predict physical activity during Covid-19 [30]. Passive sensing can also support prescriptive solutions to poor mental health; Sano et al. [33] developed an app that identifies risk factors and then alerts the participant with information to improve health and well-being.

Numerous studies employ a combination of passive data collection and EMA to acquire both passive data and active data, such as surveys and health assessments, from participants [35, 41]. The integration of EMA with passive sensing enables research in systematic phenomenology, human-computer interaction (HCI), user experience, marketing, and health studies. Passive data collection is also utilized in studies to validate and compare EMA data. For instance, Alder et al. [1] used both EMA and passive sensing as inputs to train machine-learning models to predict mental health symptoms, and Campbell [28] used mobile sensing and EMA to develop deep-learning models to classify college student's concerns during the COVID-19 pandemic. Domoff et al. [15] utilized passive data collection to verify EMA to quantify screen time usage in adolescents, however, they noted that passive data collection was limited due to unpredictable background runtime.

Securing permission to passively collect data often poses challenges for researchers due to the diverse authorization methods implemented by mobile devices [7]. Ensuring that users understand permission prompts and their implications has long been problematic [23]. Tan et al. [37] showed that developer-specified explanations on permission prompts resulted in users being more permissive in providing authorization to apps. More recently, Wijesekera et al. [47] explored machine-learning-assisted authorization systems, which significantly reduced the error rate of granting permissions.

Nishiyama et al. discovered that frequent EMA notifications can assist in passive data collection; however, further investigation is warranted [29]. As mobile operating systems continually update their privacy features and authorization procedures, additional efforts are necessary to establish best practices for both EMA and passive sensing [39]. Our contribution to this body of work involved examining the distinct authorization procedures employed by mobile operating systems for polling and event-driven passive data collection. In addition, we also explore persistent reminders as an alternative method to passively collect data.

3 METHODS

We developed a cross-platform framework for performing both passive sensing and EMA tasks while integrating a cloud-based data storage system. The native language of each operating system, Swift for iOS and Kotlin for Android, was used for development to provide full access to low-level sensors and usage logs. Both applications featured a home screen widget that displayed the study progress, which was used as an alternative method for polling. The main screen, EMA screen, and home screen widget are shown in Figure 2. We tested our framework in a mixed-method user study.

3.1 RQ1: Event-Driven Authorization Procedures

To understand both system-level constraints and user-centered implications of event-driven authorization procedures, we studied setup prompting, the method used by Android devices, and contextual prompting on iOS devices. In this study, we used location change events as our event-of-interest. Other events such as calls, messages, or health alerts could also be used to study event-driven device authorization. Location event permissions are obtained through a two-step process. The initial step involved requesting users to authorize access to location data while the app is in use. The second step requires users to grant access to location data while the app is in the background.

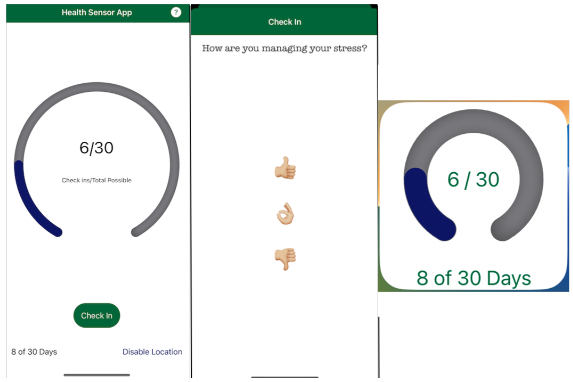

Fig. 2. Left: Main screen displaying study progress. Middle: EMA screen. Right: Home screen widget for a persistent reminder of study participation.

For setup prompting on Android devices, the process for obtaining location access is shown in Figure 3. First, the app provides a rationale for requesting location data. Next, the mobile operating system prompts the user to authorize foreground, or "while using", location access. Next, the app explains the need for background location access and provides instructions on how to grant permission. Finally, users are directed to the settings where they can grant access to background location events.

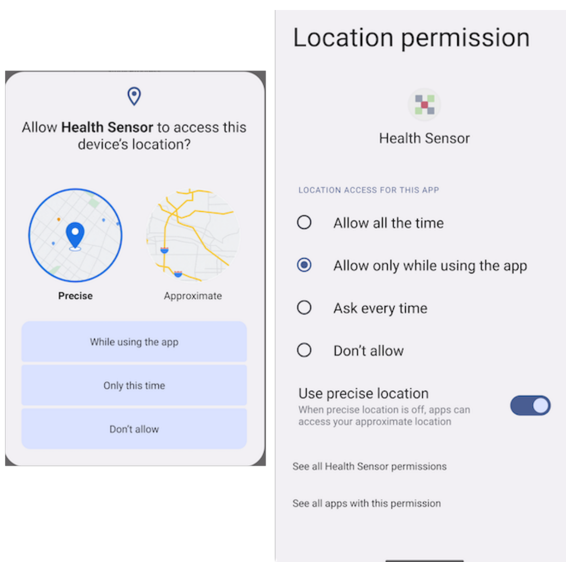

Fig. 3. Process for gaining access to background location events through setup prompting. Left: OS displays a permission prompt for foreground access. Right: Next, users are directed to the settings to allow background access ("Allow all the time").

The process for contextual prompting on iOS devices, shown in Figure 4, is to first present a rationale for requesting location access. Then, the OS prompts the user to grant foreground, or "while using", location access, similar to setup prompting. If the user grants foreground location access, then when the first location event occurs, the OS prompts the user to grant access to background location events.

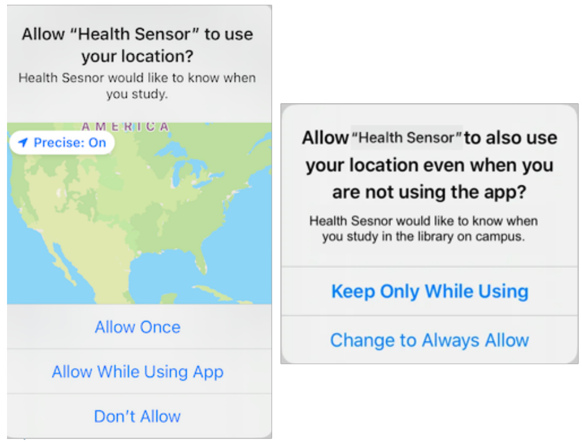

Fig. 4. Process for gaining access to background location events through contextual prompting. Left: OS displays a permission prompt for foreground access. Right: When the first location event occurs, the OS displays a second prompt for background access ("Always Allow").

3.2 RQ2: Polling

To explore the effectiveness of polling both from a systems and user level, we studied both traditional background tasks and a new approach with persistent reminders. For both iOS and Android, a machine-learning model considers both user actions, such as app usage, and system-level items such as battery status, when determining when to offer background runtime. To examine the practicality of collecting data through background tasks, we implemented a background task that requested execution time once a day. The background task logged the time and synced data to the cloud.

We also studied the feasibility of collecting data through a persistent reminder by implementing a home screen widget, shown on the right of Figure 2. Like background tasks, the widget requested a daily refresh. The OS allocates background runtime to widgets to update their content where the application can collect sensor data and usage logs.

To study user compliance with a persistent reminder, the app randomly assigned participants to a control group, that did not receive any prompts about the widget, and an experimental group where the app prompted the participant to install the app on the home screen. In addition, a separate group was given verbal instructions to install the widget.

3.3 RQ3: Notification Authorization

To study the effect that supplemental user-interface features such as notification summaries, distraction-free times, and contextual prompting have on notification delivery, we implemented a daily EMA that uses notifications to alert users to complete their assessment. Both Android and iOS devices require applications to gain user permission to send notifications. For Android devices, the authorization process uses setup prompting, where participants are presented with a request to authorize notifications before delivery is allowed.

With iOS devices, the notification authorization process uses contextual prompts. With contextual prompts, the first notification is permitted to be delivered and incorporates a prompt asking the user to grant permission for subsequent notifications, as illustrated in Figure 5.

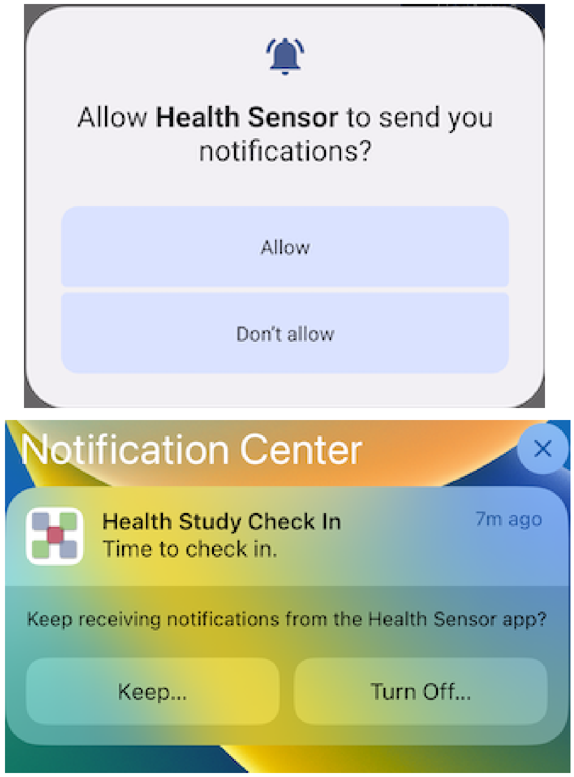

Fig. 5. Notification Permission Prompts, top: Setup Prompt, bottom: Contextual Prompt

Besides authorization procedures, Android devices and iOS devices have significantly different notification systems. On iOS, users can establish focused periods during which notifications from most apps are silently delivered and aggregated in the notification center. Notification systems on Android devices vary based on the model and can be personalized through third-party applications. Owing to these differences and the challenge of unifying notification interfaces, our study refrained from drawing quantitative conclusions based on notifications and EMA. Nevertheless, insights were derived from our qualitative analysis.

3.4 Study Procedure

We evaluated the performance of our passive sensing framework in the wild through a mixed-method user study investigating the impact of screen time and study time on college students' stress levels. The study was approved by the Institutional Research Board at (REDACTED FOR ANONYMOUS SUBMISSION).

We recruited 145 college students majoring in Computer Science, Information Technology, Business, and Science. We did not record specific demographics from the participants. The students were offered extra credit incentives to install the study application on their personal mobile devices and actively engage in the study for 30 days.

A geofence was utilized in the study to initiate event-driven data collection. Participants were informed that the application calculated study time based on their presence in the campus library and their classrooms.

The participants were allocated randomly into one of two study groups by the app: an experimental group that received a prompt from the application to install the widget on their home screen and a control group that did not. Within one of the courses we recruited from, all individuals were assigned to the widget group and were verbally instructed to install the widget during their class. Help was also offered to participants who were unfamiliar with how to install a widget on their devices.

Upon completion of the study, a follow-up survey was sent to all participants which was completed by 48 of the participants. We intended to gather qualitative insights regarding the underlying interface shifts that led to the differing success rates across passive sensing authorization paradigms. The participants were asked about: (1) what reminded them to complete their assessment, (2) their general thoughts on granting background location access, (3) whether they installed the home screen widget and their corresponding thoughts about the widget, and (4) whether they had any suggestions to improve the app.

3.5 Data Cleaning and Analysis

We initially logged 178 (126 iOS, 52 Android) users who installed the app and consented to participate in the study. We removed all users from our final analysis who completed less than 5 EMA tasks. This yielded a total of 145 participants with 105 iOS devices and 40 Android devices included in our final analysis.

Users who authorized background location access were identified as those who had 10 or more location events logged since some location events were generated when the user launched the app while inside the geofence. Testing before the study showed that, once background access was authorized, both Android and iOS devices reported the same number of events, given the same movement pattern.

To conduct our qualitative analysis, we coded responses to the survey based on general patterns that emerged in at least 5 responses. Through iterative groups and labeling, some common themes emerged as described below.

4 RESULTS

145 participants completed the study, comprising 105 iOS users and 40 Android users. Each participant had an equal chance of being assigned to the widget or control group by the app. After eliminating users who did not complete the study, there were 52 individuals (34 iOS and 18 Android) who were randomly assigned to the widget group, where the application prompted them to install the widget. 82 participants were randomly assigned to the control group, which did not receive the prompts. In addition, 11 participants (8 iOS, 3 Android) received explicit verbal instructions to install the widget on their home screens.

4.1 RQ1: Event-Driven Passive Data Collection

4.1.1 Quantitative Results. A total of 27.5% of the Android users and 48.6% of the iOS users enabled background location permission. A significant difference between contextual prompting and setup prompting was observed (X^2=4.42, df=1, $p < 0.035$). Thus, contextual prompting was 55.46% more effective than setup prompting in gaining authorization for background location events. Figure 6 illustrates the total number of location events collected, categorized by authorization method.

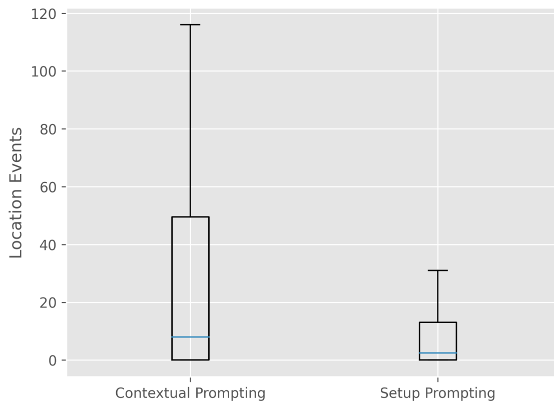

Fig. 6. Location Events by authorization method. Each location event represents an opportunity to passively collect data.

4.1.2 Qualitative Observations. Most surveyed participants did not have any concerns about enabling location sharing. Only one participant, P10 (iOS), expressed apprehension about sharing their location, stating "*I was afraid it could track me.*" However, most users expressed a willingness to share their location to contribute to the study's objectives and thought they granted the needed authorization to access background location events. P46 (Android) exemplified this sentiment, stating, "*Yes, if it helps the study then I don't really mind if it has my location.*" User's willingness to share location data for non-research reasons will most likely vary, so we emphasize that these results only apply to the context of research studies.

4.2 RQ2: Passive Data Collection through Polling

4.2.1 Quantitative Results. An ANOVA test revealed a significant difference in the total number of background sessions between iOS devices and Android devices ($F(1,144) = 37.8$, $p < 0.001$), showing that Android devices were more permissive in granting background runtime. The majority of iOS devices completed fewer than 10 background sessions, and many devices did not execute a single background task. In addition, almost half of the Android devices did not complete a daily background task each day. This demonstrates the difficulty researchers have in consistently gaining background runtime. Figure 7 provides a visual representation of the total number of background sessions conducted by each participant's device.

A total of 19 participants installed the home screen widget, with 17 being iOS users and 2 being Android users, as shown in Table 1. Of the 11 iOS users in the class who received verbal instructions to install the widget, 5 complied accordingly (45.5%), which is comprised of 4 iOS users and 1 Android user.

Group	iOS (Installed/Total)	Android (Installed/Total)
Control	1/63 (1.6%)	0/19 (0%)
Widget	12/34 (35.3%)	1/18 (5.6%)
Verbal Instructions	4/8 (50%)	1/3 (33.3%)

Table 1. Users that installed the widget

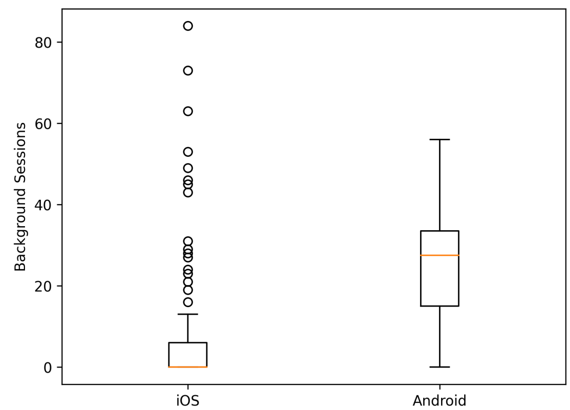

Fig. 7. Total number of background tasks completed, by device.

Of the participants who installed the home screen widget on iOS devices, the widget was refreshed on average 64.5 times, with a standard deviation of 13.7, as depicted in Figure 8. The two Android devices refreshed 23 and 42 times. Focusing solely on iOS widget users, their devices exhibited an M=7.24 (SD=13.24) background tasks and M=64.53 (SD = 13.76) widget refreshes throughout the study, which represents a 795.8% increase. An ANOVA test ($F(1,18) = 153.52$, $p < 0.001$) yielded a significant difference in the number of widget refreshes compared to the number of background tasks.

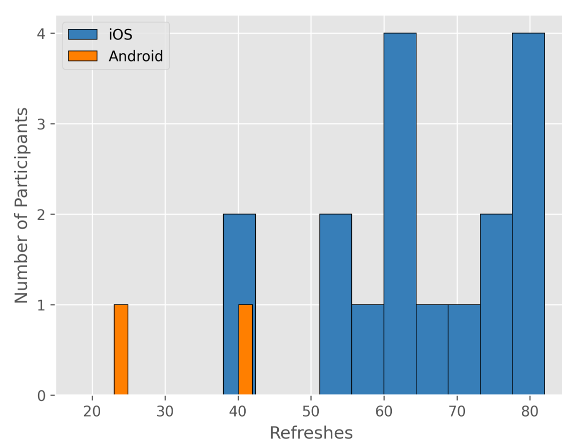

Fig. 8. Histogram of widget refreshes. Each widget refresh presents an opportunity to passively collect data.

4.2.2 *Qualitative Observations.* Several participants indicated that the widget aided them in remembering to complete their daily assessment, as expressed by P15, P24, and P45. Specifically, P37 (iOS) stated, "*The widget helped me to remember to check in every day and what the study was about.*" However, a few users opted not to utilize the widget due to concerns about modifying their home screen. P22 (iOS) mentioned, "*My home screen was already full and organized*," while P35 (Android) stated, "*I just don't want to change my home screen.*" Additionally, some users cited unfamiliarity with widget usage as the reason for not using it, as mentioned by P26 (Android) and P42 (iOS).

The majority of surveyed participants expressed no reservations about using the widget. iOS users may exhibit a greater willingness to add the widget to their home screen because iOS allows for widget stacking, a feature that is not found in most Android devices. This functionality enables the rotation of different widgets within a stack, eliminating the need to reorganize the home screen to install the widget.

4.3 RQ3: EMA Compliance and Notifications

4.3.1 Quantitative Results. We observed no statistically significant difference in the number of completed assessments when examining device-specific differences (see Figure 9) as shown by an ANOVA test ($F(1,144) = 0.0227$, $p = 0.88$) or when comparing the widget group to the control group ($F(1,144) = 1.33$, $p = 0.268$).

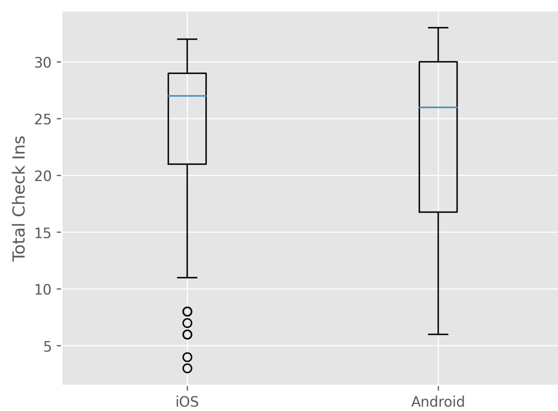

Fig. 9. EMA completed by device. There was no statistical difference between devices in the total number of assessments completed.

4.3.2 Qualitative Observations. Notifications played a role in reminding participants to complete the EMA task, particularly for Android users (P13, P16, P17, P19). Despite some iOS users (P7, P9, P18) stating that the notifications help them remember to complete the task, a majority of the iOS users surveyed said the notifications did not appear consistently. Some participants (P3, P5, P29, P23, P43, P47) even set their own alarms or reminders: *"Notifications weren't working, so I had an event in my calendar to remind me."* (P32)

5 DISCUSSION AND DESIGN GUIDELINES

Drawing upon both our quantitative and qualitative analyses, we present design principles to enhance the success of mobile passive-sensing research studies:

For polling, implement a persistent reminder that can be used both as a means of collecting data and as a reminder to complete EMA tasks. Collecting data through background tasks presents significant challenges due to the use of an implied consent model for granting background runtime by mobile operating systems. Although Android devices are generally more permissive, there are too many factors under consideration to consistently guarantee background runtime. However, we found that using a persistent reminder as a secondary means to poll for data yielded more successful data retrieval. The installation of a persistent reminder on the user's home screen signifies to the operating system the intent to allocate the necessary resources to maintain the reminder's regular updates, which includes the collection of data.

Our survey revealed that many users did not comply with installing the widget when prompted by the app, especially among Android users. Verbal prompts to install the widget did help improve compliance on both devices. Our qualitative analysis indicated that there was some resistance to installing the widget because users did not want to change their home screen layout, and some users were unfamiliar with how to install widgets on their devices. We postulate that integrating widget installation as a prerequisite for earning study incentives could potentially enhance compliance rates.

For event-driven data collection, use carefully designed contextual prompts. Our qualitative analysis revealed that most participants were willing to authorize background location access for research studies and were under the impression that they enabled background location access. However, our quantitative results revealed that many participants did not authorize background location access. Contextual prompting on iOS was 55.46% more effective in gaining authorization for background location events. This coincides with previous work that shows providing additional context helps users make better privacy decisions [5, 6, 46].

However, even with contextual prompts, less than half of the participants authorized access to background location events. To improve contextual prompt accuracy Wijesekera et al. [47] developed a machine learning-based, contextual-aware permission model that improves the permission accuracy rate above context prompts, which should be considered if it becomes available on consumer mobile devices. However, the current model suggests providing "generic but well-formed data" when an app is denied access. In the context of health sensing apps, this can skew results and create adverse health interventions. We recommend that such systems be designed to communicate to the app about denials and that apps be constructed to handle denials to ensure data fidelity.

Setup prompts should be used for notifications. Through our qualitative analysis, we discovered that many iOS users did not consistently receive notifications to remind them to check in. Current iOS notification systems include distraction-free times, notification summaries, "send now, never alert" notifications, etc. described by Li et al. [26]. These additions to notification systems require additional effort to ensure that participants receive alerts to complete EMA tasks.

Sending notifications require authorization. Normally contextual prompts tend to work more effectively, as described above, but with the introduction of "send now, never alert" notifications or notification deferral, contextual prompts lose their context. If users never receive a context prompt, authorization could be incorrectly denied. Only users that searched through their notification summaries would be able to find the contextual prompt and authorize future notifications. As intelligent notification systems [4, 26] become available on consumer mobile devices, EMA study apps will need to evolve to ensure proper notification delivery. Intelligent notification systems should also include methods to allow users to ensure EMA apps receive authorization delivery notifications.

5.1 Study Limitations and Future Work

There are other differences between iOS and Android devices and users of those devices than what we tested for, and those could be confounding factors. In addition, we studied passive sensing for research studies and the results do not necessarily apply to passive sensing in other contexts, such as for commercial purposes. Like many HCI studies, this study involved college students, and while they represented diverse academic disciplines and countries of origin, more work could be conducted to verify that the results apply to a wider population.

For persistent reminders, only two Android users installed the widget, with one device refreshing more than daily, and the other device refreshing 23 out of the 30 days. These preliminary results are promising, however further work is needed to ensure that the results are consistent with a larger sample size.

Because we used the mobile devices' default authorization procedures, we did not test changes to the prompts. More work is needed to refine contextual prompts for background events to ensure willing participants correctly authorize event-driven data collection. Future work could include testing various contextual prompts for background events to optimize acceptance rates among willing participants. With notification systems continuing to add features to minimize distractions, more work is required to ensure that willing participants can correctly authorize notification delivery, which is an important feature for EMA applications.

6 CONCLUSION

We developed a passive sensing framework using both iOS and Android devices, which we subsequently used to assess the authorization process of passively collecting data through both polling and event-driven collection methods. We tested our sensing framework in a study with 145 college students that purported to study how screentime and study habits affect stress levels in college students. Through both quantitative and qualitative analysis, we found that for event-driven data collection, contextual prompts work better for gaining authorization for background events, but further work is needed to design prompts that correctly convey the privacy implications of granting background access.

For polling, we found that gaining authorization to run background tasks continues to be a challenge due to the implied consent model mobile operating systems use to preserve battery life. We explored using a home screen widget as a persistent reminder. The widget not only helped participants remember to complete their EMA but also provided an additional opportunity to poll for sensor data during refreshes.

As a result of our analysis, we provide design guidelines that can be incorporated into passive sensing frameworks to ensure that participants willingly authorize passive data collection. The design guidelines are meant to help researchers develop passive sensing frameworks that cooperate with mobile OS's goals to protect user privacy and optimize battery performance while still enabling the flexibility required for longitudinal passive sensing studies.

REFERENCES

[1] Daniel A Adler, Fei Wang, David C Mohr, and Tanzeem Choudhury. 2022. Machine learning for passive mental health symptom prediction: Generalization across different longitudinal mobile sensing studies. *Plos one* 17, 4 (2022), e0266516.

[2] Joost Asselbergs, Jeroen Ruwaard, Michal Ejdys, Niels Schrader, Marit Sijbrandij, Heleen Riper, et al. 2016. Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. *Journal of medical Internet research* 18, 3 (2016), e5505.

[3] Karim Assi, Lakmal Meegahapola, William Droz, Peter Kun, Amalia De Götzen, Miriam Bidoglia, Sally Stares, George Gaskell, Altangerel Chagnaa, Amarsanaa Ganbold, et al. 2023. Complex Daily Activities, Country-Level Diversity, and Smartphone Sensing: A Study in Denmark, Italy, Mongolia, Paraguay, and UK. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–23.

[4] Jonas Auda, Dominik Weber, Alexandra Voit, and Stefan Schneegass. 2018. Understanding user preferences towards rule-based notification deferral. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

[5] Adam Barth, Anupam Datta, John C Mitchell, and Helen Nissenbaum. 2006. Privacy and contextual integrity: Framework and applications. In *2006 IEEE symposium on security and privacy (S&P'06)*. IEEE, 15–pp.

[6] Bram Bonné, Sai Teja Peddinti, Igor Bilogrevic, and Nina Taft. 2017. Exploring decision making with {Android's} runtime permission dialogs using in-context surveys. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. 195–210.

[7] Tjeerd W Boonstra, Jennifer Nicholas, Quincy JJ Wong, Frances Shaw, Samuel Townsend, and Helen Christensen. 2018. Using mobile phone sensor technology for mental health research: integrated analysis to identify hidden challenges and potential solutions. *Journal of medical Internet research* 20, 7 (2018), e10131.

[8] Larry Chan, Vedant Das Swain, Christina Kelley, Kaya de Barbaro, Gregory D Abowd, and Lauren Wilcox. 2018. Students' experiences with ecological momentary assessment tools to report on emotional well-being. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–20.

[9] Sydney Charitos, Amberly Brigden, and Jon Bird. 2023. DIY Ecological Momentary Assessment. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–4.

[10] 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. 1–14.

[11] Vedant Das Swain, Lan Gao, William A Wood, Srikruthi C Matli, Gregory D Abowd, and Munmun De Choudhury. 2023. Algorithmic Power or Punishment: Information Worker Perspectives on Passive Sensing Enabled AI Phenotyping of Performance and Wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[12] Kristen Dergousoff and Regan L Mandryk. 2015. Mobile gamification for crowdsourcing data collection: Leveraging the freemium model. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1065–1074.

[13] Kevin Doherty, Andreas Balaskas, and Gavin Doherty. 2020. The design of ecological momentary assessment technologies. *Interacting with Computers* 32, 1 (2020), 257–278.

[14] Kevin Doherty, José Marcano-Belisario, Martin Cohn, Nikolaos Mastellos, Cecily Morrison, Josip Car, and Gavin Doherty. 2019. Engagement with mental health screening on mobile devices: Results from an antenatal feasibility study. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.

[15] Sarah E Domoff, Claire Ann Banga, Aubrey L Borgen, Ryan P Foley, Chelsea Robinson, Katie Avery, and Douglas A Gentile. 2021. Use of passive sensing to quantify adolescent mobile device usage: Feasibility, acceptability, and preliminary validation of the eMoodie application. *Human Behavior and Emerging Technologies* 3, 1 (2021), 63–74.

[16] Afsaneh Doryab, Daniella K Villalba, Prerna Chikersal, Janine M Dutcher, Michael Tumminia, Xinwen Liu, Sheldon Cohen, Kasey Creswell, Jennifer Mankoff, John D Creswell, et al. 2019. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and fitbit data. *JMIR mHealth and uHealth* 7, 7 (2019), e13209.

[17] Hui Gao, Chi Harold Liu, Jian Tang, Dejun Yang, Pan Hui, and Wendong Wang. 2018. Online quality-aware incentive mechanism for mobile crowd sensing with extra bonus. *IEEE Transactions on Mobile Computing* 18, 11 (2018), 2589–2603.

[18] Yu-Ning Huang, Siyan Zhao, Michael L Rivera, Jason I Hong, and Robert E Kraut. 2021. Predicting well-being using short ecological momentary audio recordings. In *Extended abstracts of the 2021 CHI conference on human factors in computing systems*. 1–7.

[19] Stephen S Intille, Emmanuel Munguia Tapia, John Rondoni, Jennifer Beaudin, Chuck Kukla, Sitij Agarwal, Ling Bao, and Kent Larson. 2003. Tools for studying behavior and technology in natural settings. In *UbiComp 2003: Ubiquitous Computing: 5th International Conference, Seattle, WA, USA, October 12-15, 2003. Proceedings 5*. Springer, 157–174.

[20] Luis G Jaimes, Idalides J Vergara-Laurens, and Andrew Raij. 2015. A survey of incentive techniques for mobile crowd sensing. *IEEE Internet of Things Journal* 2, 5 (2015), 370–380.

[21] Niels Jongs, Raj Jagesar, Neeltje EM van Haren, Brenda WJH Penninx, Lianne Reus, Pieter J Visser, Nic JA van der Wee, Ina M Koning, Celso Arango, Iris EC Sommer, et al. 2020. A framework for assessing neuropsychiatric phenotypes by using smartphone-based location data. *Translational psychiatry* 10, 1 (2020), 1–10.

[22] Soowon Kang, Cheul Young Park, Auk Kim, Narae Cha, and Uichin Lee. 2022. Understanding Emotion Changes in Mobile Experience Sampling. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.

[23] Patrick Gage Kelley, Sunny Consolvo, Lorrie Faith Cranor, Jaeyeon Jung, Norman Sadeh, and David Wetherall. 2012. A conundrum of permissions: installing applications on an android smartphone. In *Financial Cryptography and Data Security: FC 2012 Workshops, USEC and WECSR 2012, Kralendijk, Bonaire, March 2, 2012, Revised Selected Papers 16*. Springer, 68–79.

[24] Aaron Kline, Catalin Voss, Peter Washington, Nick Haber, Hessey Schwartz, Qandeel Tariq, Terry Winograd, Carl Feinstein, and Dennis P Wall. 2019. Superpower glass. *GetMobile: Mobile Computing and Communications* 23, 2 (2019), 35–38.

[25] Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T Campbell. 2010. A survey of mobile phone sensing. *IEEE Communications magazine* 48, 9 (2010), 140–150.

[26] Tianshi Li, Julia Katherine Haines, Miguel Flores Ruiz De Eguino, Jason I Hong, and Jeffrey Nichols. 2023. Alert Now or Never: Understanding and Predicting Notification Preferences of Smartphone Users. *ACM Transactions on Computer-Human Interaction* 29, 5 (2023), 1–33.

[27] Robert LiKamWa, Yunxin Liu, Nicholas D Lane, and Lin Zhong. 2013. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. 389–402.

[28] Subigya Nepal, Weichen Wang, Vlado Vojdanovski, Jeremy F Huckins, Alex Dasilva, Meghan Meyer, and Andrew Campbell. 2022. COVID student study: A year in the life of college students during the COVID-19 pandemic through the lens of mobile phone sensing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–19.

[29] Yuuki Nishiyama, Denzil Ferreira, Yusaku Eigen, Wataru Sasaki, Tadashi Okoshi, Jin Nakazawa, Anind K Dey, and Kaoru Sezaki. 2020. IOS crowd–sensing won't hurt a bit!: AWARE framework and sustainable study guideline for iOS platform. In *Distributed, Ambient and Pervasive Interactions: 8th International Conference, DAPI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22*. Springer, 223–243.

[30] Agnes Norbury, Shelley H Liu, Juan José Campaña-Montes, Lorena Romero-Medrano, María Luisa Barrigón, Emma Smith, Antonio Artés-Rodríguez, Enrique Baca-Garcia, and M Mercedes Perez-Rodriguez. 2021. Social media and smartphone app use predicts maintenance of physical activity during Covid-19 enforced isolation in psychiatric outpatients. *Molecular psychiatry* 26, 8 (2021), 3920–3930.

[31] Mashfiqui Rabbi, Shahid Ali, Tanzeem Choudhury, and Ethan Berke. 2011. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In *Proceedings of the 13th international conference on Ubiquitous computing*. 385–394.

[32] John Rooksby, Alistair Morrison, and Dave Murray-Rust. 2019. Student perspectives on digital phenotyping: The acceptability of using smartphone data to assess mental health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[33] Akane Sano, Sara Taylor, Andrew W McHill, Andrew JK Phillips, Laura K Barger, Elizabeth Klerman, and Rosalind Picard. 2018. Identifying Objective Physiological Markers and Modifiable Behaviors for Self-Reported Stress and Mental Health Status Using Wearable Sensors and Mobile Phones: Observational Study. *J Med Internet Res* 20, 6 (08 Jun 2018), e210. https://doi.org/10.2196/jmir.9410

[34] Saul Shiffman, Arthur A Stone, and Michael R Hufford. 2008. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* 4 (2008), 1–32.

[35] Dimitris Spathis, Sandra Servia-Rodriguez, Katayoun Farrahi, Cecilia Mascolo, and Jason Rentfrow. 2019. Passive mobile sensing and psychological traits for large scale mood prediction. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 272–281.

[36] Arthur A Stone, Stefan Schneider, and Joshua M Smyth. 2023. Evaluation of Pressing Issues in Ecological Momentary Assessment. *Annual Review of Clinical Psychology* 19 (2023).

[37] Joshua Tan, Khanh Nguyen, Michael Theodorides, Heidi Negrón-Arroyo, Christopher Thompson, Serge Egelman, and David Wagner. 2014. The effect of developer-specified explanations for permission requests on smartphone user behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 91–100.

[38] Yonatan Vaizman, Katherine Ellis, Gert Lanckriet, and Nadir Weibel. 2018. Extrasensory app: Data collection in-the-wild with rich user interface to self-report behavior. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

[39] 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.

[40] Catalin Voss, Jessey Schwartz, Jena Daniels, Aaron Kline, Nick Haber, Peter Washington, Qandeel Tariq, Thomas N Robinson, Manisha Desai, Jennifer M Phillips, et al. 2019. Effect of wearable digital intervention for improving socialization in children with autism spectrum disorder: a randomized clinical trial. *JAMA pediatrics* 173, 5 (2019), 446–454.

[41] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 3–14.

[42] Weichen Wang, Shayan Mirjafari, Gabriella Harari, Dror Ben-Zeev, Rachel Brian, Tanzeem Choudhury, Marta Hauser, John Kane, Kizito Masaba, Subigya Nepal, et al. 2020. Social sensing: assessing social functioning of patients living with schizophrenia using mobile phone sensing. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–15.

[43] Peter Washington, Aaron Kline, Onur Cezmi Mutlu, Emilie Leblanc, Cathy Hou, Nate Stockham, Kelley Paskov, Brianna Chrisman, and Dennis Wall. 2021. Activity recognition with moving cameras and few training examples: applications for detection of autism-related headbanging. In *Extended abstracts of the 2021 CHI conference on human factors in computing systems*. 1–7.

[44] Peter Washington, Catalin Voss, Nick Haber, Serena Tanaka, Jena Daniels, Carl Feinstein, Terry Winograd, and Dennis Wall. 2016. A wearable social interaction aid for children with autism. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2348–2354.

[45] Peter Washington, Catalin Voss, Aaron Kline, Nick Haber, Jena Daniels, Azar Fazel, Titas De, Carl Feinstein, Terry Winograd, and Dennis Wall. 2017. SuperpowerGlass: a wearable aid for the at-home therapy of children with autism. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–22.

[46] Primal Wijesekera, Arjun Baokar, Ashkan Hosseini, Serge Egelman, David Wagner, and Konstantin Beznosov. 2015. Android permissions remystified: A field study on contextual integrity. In *24th USENIX Security Symposium (USENIX Security 15)*. 499–514.

[47] Primal Wijesekera, Joel Reardon, Irwin Reyes, Lynn Tsai, Jung-Wei Chen, Nathan Good, David Wagner, Konstantin Beznosov, and Serge Egelman. 2018. Contextualizing privacy decisions for better prediction (and protection). In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[48] 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*. 1735–1740.

Received 17 Sept 2023; revised 12 March 2023; accepted 5 June 2023