Check for updates

Pinning, Sorting, and Categorizing Notifications: A Mixed-methods Usage and Experience Study of Mobile Notification-management Features

YONG-HAN LIN, National Yang Ming Chiao Tung University, Taiwan
LI-TING SU, National Yang Ming Chiao Tung University, Taiwan
UEI-DAR CHEN, National Yang Ming Chiao Tung University, Taiwan
YI-CHI LEE, National Yang Ming Chiao Tung University, Taiwan
PENG-JUI WANG, National Yang Ming Chiao Tung University, Taiwan
YUNG-JU CHANG*, National Yang Ming Chiao Tung University, Taiwan

As smartphone notifications increase, so does the effort required to handle them effectively. Previous research has proposed various notification management features, but empirical evidence regarding their efficacy remains sparse. In response, we developed a notification management application incorporating features derived from prior studies, including both automatic and manual sorting, categorization, and manual pinning. Utilizing a mixed-methods approach, we explored how users interact with these features in their daily routines, with the aim to identify the underlying needs driving their usage. The results indicate that pinning was the most valued feature, serving diverse purposes such as deferring notifications, ensuring quick and constant access to information, preventing accidental deletions, and providing visual reminders. Conversely, manual categorization was underutilized, with participants relying on automated categories for notification access. Moreover, participants expressed a desire for automatic features to process and organize notifications based on topic and personalize them through user input. They also expected automatic sorting to adapt more dynamically to user contexts.

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

Additional Key Words and Phrases: smartphone notifications; notification management; pinning; sorting; categorizing

ACM Reference Format:
Yong-Han Lin, Li-Ting Su, Uei-Dar Chen, Yi-Chi Lee, Peng-Jui Wang, and Yung-Ju Chang. 2024. Pinning, Sorting, and Categorizing Notifications: A Mixed-methods Usage and Experience Study of Mobile Notification-management Features. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 115 (September 2024), 27 pages. https://doi.org/10.1145/3678579

*Corresponding author.

Authors' addresses: Yong-Han Lin, alberta.cs07@nycu.edu.tw, National Yang Ming Chiao Tung University, 1001 University Road, Hsinchu, Taiwan, Taiwan, 30010; Li-Ting Su, liitiisu.cs10@nycu.edu.tw, National Yang Ming Chiao Tung University, 1001 University Road, Hsinchu, Taiwan, Taiwan, 30010; Uei-Dar Chen, udchen.cs08@nycu.edu.tw, National Yang Ming Chiao Tung University, 1001 University Road, Hsinchu, Taiwan, Taiwan, 30010; Yi-Chi Lee, yichi170.cs08@nycu.edu.tw, National Yang Ming Chiao Tung University, 1001 University Road, Hsinchu, Taiwan, Taiwan, 30010; Peng-Jui Wang, james900425.cs08@nycu.edu.tw, National Yang Ming Chiao Tung University, 1001 University Road, Hsinchu, Taiwan, Taiwan, 30010; Yung-Ju Chang, armuro@cs.nctu.edu.tw, National Yang Ming Chiao Tung University, 1001 University Road, Hsinchu, Taiwan, Taiwan, 30010.

2474-9567/2024/9-ART115
https://doi.org/10.1145/3678579

1 INTRODUCTION

The proliferation of mobile apps is generating an increasingly diverse array of notifications. This not only elevates the potential for interruptions but also demands that smartphone interfaces and functionalities evolve to help users efficiently identify relevant notifications based on their immediate needs [11, 28, 74] and handle notifications that serve different roles [43]. However, while the disruptive nature of notifications has long been a primary focus of research [13, 31, 35], typical solutions have involved either automatically identifying or allowing users to designate suitable times for notification delivery [1, 9, 25, 49, 52, 54, 56, 57, 73].

In recent years, as the number and variety of notifications have grown, several studies have begun to investigate how mobile users manage and organize notifications [4, 11, 41, 43, 54, 57, 59, 70, 73, 74], with some exploring the usage of specific features [41, 54, 73]. These studies have highlighted diverse user actions such as immediate reading and responding [41, 43, 57, 59], saving for later [43], snoozing [41, 54, 73], leaving unread [4, 11, 70, 74], swiping off [41, 43, 70], or bulk removing without reading [41, 70, 74], reflecting the varying relevance, urgency, and functionality of notifications. However, these findings also underscore that current notification interfaces are still inadequate in supporting the varied roles of notifications, such as efficiently locating relevant notifications for immediate needs [11, 43], deferring notifications to an unspecified later time [73], and organizing them systematically [43]. In response to these limitations, new features have been proposed to enhance support for these needs, including improved sorting [11, 43], categorizing [28, 74, 75], and pinning for deferred access [43], assuming that these features could support users in handling the increasing array of notifications.

However, relatively few of the features that researchers have proposed to meet these divergent user needs have been empirically evaluated and investigated. Thus, it remains unclear how users would utilize them to interact with a diverse array of notifications, whether these features can support users in effectively and efficiently handling notifications, what shortcomings exist, and how they can be improved to better meet user needs. We regard these questions as crucial, both for industry practitioners making development decisions, and for researchers deciding which proposed features merit further investigation.

Therefore, this study's aim is to evaluate the three main types of features that prior research has proposed would aid users' smartphone-notification management: i.e., sorting [11, 43], categorizing [28, 74, 75], and pinning [43]. To help us achieve it, we developed NotiManager, an Android research application that encompasses all three of these features. Additionally, because of prior research claims about the advantages of both manual control (i.e., that it confers greater user autonomy) and automatic features (i.e., convenience) [44], our app included both manual and automatic approaches to sorting and categorizing, as a means of clarifying whether one of them – or a combination of both – was more beneficial.

Our research questions are:

- RQ1: How would users utilize sorting, categorization, and pinning in a notification interface when these features are available?
- RQ2: How do users perceive and experience these three features in managing their notifications?

To address these questions, we recruited 30 Android users to interact with all three features in their day-to-day lives. Our research app was developed in three distinct modes: Manual, where users manually controlled all features; Auto, which automated all features except for pinning; and Hybrid, which combined elements of both manual and automatic controls. Using a mixed-method approach, we examined how users interacted with, perceived, and experienced each mode over a three-week period. This paper reports our findings, including their usage, perceptions, experiences, as well as underlying needs associated with notification management and proposes enhancements to better support these needs in future notification systems. Specifically, this study highlights four key contributions to the literature on notification management:

- It identifies pinning as a highly valued and utilized feature for its versatility in notification management, including deferring notifications, ensuring quick and constant access to information, temporarily preserving notifications to prevent accidental deletions, and providing visual reminders.
- It reveals a potential emerging workflow in notification management when pinning is available—pin to temporarily preserve, bulk remove others, and then assess the pinned notifications.
- It uncovers how users make use of manual sorting to organize notifications and notes the low usage of manual categorization for organizing and accessing notifications.
- It shows users' expectations for the two automatic features to process and organize notifications based on topics and to take user input to better align with their preferences and needs.

2 RELATED WORK

The domain of notification-related research within HCI is broad, but the majority of work in it focuses on determining the optimal timing for notification delivery to minimize disruption [14, 26, 54, 60] or enhance user engagement [2, 24, 55, 56], or controlling notification alerts to strike a balance between user awareness and potential disturbances [13, 32, 45, 47, 57]. This paper does not aim to provide an exhaustive review of such studies, as they primarily relate to users' initial reactions to notifications. Instead, our focus is on the smaller body of work on how users manage and engage with notifications, particularly within their devices' notification centers/drawers [45, 74]. The latter subject encompasses what many researchers refer to as notification attendance and management.

Research in notification attendance and management on mobile devices has thoroughly examined user engagement, dividing behaviors into attending and responding [13, 21, 37, 38, 45, 57, 58, 77]. Attending ranges from brief examinations of notifications to focused and committed engagement with them [21]. Quick interactions are typically measured by users unlocking their phones and viewing notifications [13, 45, 57], whereas longer engagement is measured by actions like tapping or dismissing them [13].

Some studies have delved into the dynamics of attention, defining the gap between initial noticing and engagement as "decision time", and the gap between a notification's arrival and user interaction as "reaction time" [12, 30, 62]. However, some scholars consider phone-unlocking and notification-tapping to attendance [13], recognizing the nuanced and sometimes simultaneous nature of these actions. Such distinctions have led to the use of the specific terms "glancing" for quick assessment and "reading" [10, 66, 68] for committed interaction, alongside other such terms including "seeing" [59], "focusing" [68], and "engaging" [10, 67]. As a result, the dialogue surrounding the terminology for notification engagement has been complex but has yielded two essential understandings. First, the terms – and the interactions they describe – underline that user engagement with notifications is multi-stage [10, 66]. And, despite the diversity of the terminology used, all these interactions are encompassed within the broad concept of *notification management*.

The journey from noticing a notification to taking subsequent actions may entail various cognitive decisions [3, 17, 39]. As such, the literature has repeatedly called for notification systems to be endowed with new features that will support better-informed user choices when dealing with a large volume of notifications. Among these, instant messaging (IM) notifications stand out, garnering the most immediate responses [4, 57, 59, 61, 75]. Yet, it's worth noting that previous studies have shown users don't always promptly respond to instant messages [17, 21, 39]. At times, they delay attending to these messages for various reasons. This tendency to defer responses has necessitated the introduction of supportive features in notification systems, such as a deferral feature, which provides a subsequent prompt as a reminder to users [63, 73]. After all, conventional deferral techniques, whether by setting a specific duration or choosing a designated time, come with their own set of challenges, as delineated by Weber et al. [73]. Moving beyond just snoozing notifications, prior research has proposed an approach that allowed users to manually reposition notifications [43]. This stemmed from their study results that their participants hoped certain notifications to be placed in specific spots as reminders. The study further recommended automation in notification sequencing to mirror the user's preferred arrangement and proposed allowing users to "pin" specific notifications for later attention. Furthermore, some research has explored system-level features that enable users to categorize their notifications, akin to the labeling and filtering functionalities available in many email systems [28, 43, 74]. Given the complex and rich information present in both notifications and emails, a well-designed information architecture presumably provides a more organized and layered user experience.

Collectively, previous research on notification management has recommended features like pinning, sorting, and categorizing to assist mobile users. While these studies discussed the potential uses and benefits of the features, how they are actually utilized remains underexplored. We address this gap in the literature by providing empirical evidence on how smartphone users engage with these features in their daily lives and whether these features are genuinely helpful to users. Weber et al. [73] have touched on why users might want to defer notifications, but whether the reasons for pinning notifications align with those for deferring them is also unknown. Our study uses empirical research to delve into the specific reasons that drive both the use and non-use of these features, providing valuable insights for designers and developers on the potential advantages of incorporating these elements into notification systems. Moreover, we identify areas for further improvement, thereby contributing a grounded understanding of user needs and preferences around these features that can inform future notification system design.

3 NOTIMANAGER

NotiManager is an Android app developed to support users in managing notifications. It replaces the original notification drawer as the sole interface for users to interact with notifications. Beyond the basic functionalities in the existing notification drawer (i.e. clicking and deleting notifications), NotiManager offers additional features for managing notifications, including sorting, categorizing, and pinning. It monitors and captures incoming notifications through the utilization of Android Notification Listener Service API[1]. Installing NotiManager on the phone will not affect notifications popping up at the top and appearing in the notification drawer when they initially arrive. After five seconds—the average duration we found for how long a current notification pops up and alerts or vibrates—notifications, with the exception of ongoing ones (e.g. navigation, media playback notifications), will seamlessly integrated from the native notification drawer into the NotiManager interface. We've retained these ongoing notifications in the original drawer to ensure users can interact with them just as usual. For other notifications, once they disappear from the notification drawer, users need to locate and tap the NotiManager app on their phone screen to access them. The subsequent sections delve into a detailed exploration of features and interfaces in NotiManager.

3.1 Notification Management Features

We incorporated the three features guided by recommendations from existing notification research. The following section provides an introduction to these features.

(1) **Sorting:** The sorting feature reorders users' notifications, with automatic and manual methods. In automatic sorting, the complete list of notifications is transmitted to a remote server designed to sort notifications whenever a new notification arrives. This model, trained with a pre-trained BERT model[2], is based on a dataset previously collected by Lin et al. [43], where participants in their study sorted six sampled notifications displayed on the smartphone screen each time they responded to an experience sampling method (ESM) questionnaire. The model employs various features for sorting, including the notification's title and content, the originating application, the predefined category of the application, as well as user

[1] Notification Listener Service API: https://developer.android.com/reference/android/service/notification/NotificationListenerService
[2] BERT model documentation: https://huggingface.co/docs/transformers/model_doc/bert

context factors like activities[3] detected by the mobile sensors and the time when notifications were viewed. The model begins by assessing the probability that each notification belongs to one of six positions, as determined from previous study samples. To accommodate real-world scenarios where the number of notifications may differ, we calculate a weighted importance score for each notification, which determines its eventual placement in the user's view. The sorted results are then sent back from the remote server and replaced with the original order, presenting a newly arranged display. It is essential to note that since the primary goal of our automatic sorting feature was to assess user experiences and perceptions, the main intent was not to predict notification orders with the precision of leading models, but rather to utilize the model as an exploratory tool that prompts participants to identify potential areas for enhancement in future automatic sorting. Manual sorting, on the other hand, provides users with the ability to manually adjust the order by dragging and dropping notifications. NotiManager memorizes these changes and preserves the positions of the manually-adjusted notifications, regardless of arrival of new notifications. This rule also applies in the Hybrid mode, a mode that combines all the features present in the Auto mode and the Manual mode. Notifications that have been manually sorted will stay in their original positions, even if the automatic sorting attempts to reposition them differently.

(2) **Categorizing:** This feature is designed to categorize notifications or process them within a category, using automatic and manual methods. Automatic categorization operates by classifying notifications based on app categories established by the Google Play Store and as suggested by prior research [61, 73, 74]. On the other hand, manual categorization permits users to classify notifications according to custom categories they define themselves or automatically generate categories by long-pressing the notification. Each notification can only be assigned to a single category, and notifications that are categorized still appear in the all-notification interface. Whether using automatic or manual categorization, users can choose to view notifications within specific categories, providing a more tailored experience compared to viewing all notifications indiscriminately. For added flexibility in the presentation of category orders, NotiManager enables users to rearrange the order of categories displayed in the category menu.

(3) **Pinning:** To enable users to anchor specific notifications within the NotiManager interface, we introduce the pinning feature. By pinning notifications, users can ensure that the pinned notification is kept and remains visible, unaffected by any actions such as clicking, swiping off, or pressing clear-all-notifications button. The pinned notification can only be removed after deselecting the pin. This feature is designed to allow users to retain notifications that they may wish to revisit or reference later within NotiManager, fulfilling a need suggested in the literature [43] for a method to remind users to revisit notifications or to undertake specific tasks related to the notifications.

3.2 User Interface

The interface varies in different modes. The all-notification interface in the Manual mode includes icons for sorting and pinning, and a long-press menu for categorizing notifications. However, the visual cue indicating whether notifications have completed automatic sorting is not shown. In the Auto mode, only the visual signal for indicating automatic sorting completion is present. As the Hybrid mode encompasses all components, we hereby introduce the complete NotiManager interface under this mode. The comprehensive NotiManager interface includes:

(1) **All-Notification Interface:** Notifications in NotiManager display the notification title, content, post time, the corresponding app name, and the app icon. NotiManager emulates the fundamental attributes of traditional smartphone notification drawers, allowing users to engage with notifications by clicking and remove notifications using swiping gestures or the clear-all-notifications button (see Figure 1c). The

[3] Google Activity Recognition API: https://developers.google.com/location-context/activity-recognition

Figure 1. All-notification interface includes: (a) Hamburger menu button for expanding the category menu; (b) Visual indication showing if the notifications are automatically sorted: green indicates sorted and red indicates unsorted; (c) Clear-all-notifications button; (d) Pinning button (before); (e) Pinning button (after); (f) Sorting button

NotiManager all-notification interface augments the standard notification drawer with a sorting (see Figure 1f) and pinning icon (see Figure 1d). The sort icon is a button for users to rearrange notifications through a drag-and-drop function, while the pin icon is a button for users to pin their notifications by pressing the icon. Long-pressing a notification allows the user to either create a new category or change the existing category of the notification in a menu. To allow users to know whether notifications have been automatically sorted, there is a visual indication (see Figure 1b) provided for user assessment. When the visual indication transitions from red to green, it signifies the progression from ongoing sorting to sorting completion.

(2) **Notification-within-Category Interface:** After clicking the hamburger menu button (see Figure 1a) on the all-notification interface, the category menu expands and shows all the categories created. Upon selecting a category name, users can view all notifications assigned to that category. By default, the user is presented with the "All notifications" category, where they can view all incoming notifications. The interface for notifications within categories is identical to that of all notifications, including the functionalities. Actions performed on notifications within categories impact the interface of all notifications.

4 FIELD STUDY

We conducted a within-subjects field study to assess the effectiveness of our notification management features and investigated our research questions. A total of 30 Android users participated in the study. More details are provided below.

4.1 Study Design

To ensure participants were familiar with viewing and attending notifications through NotiManager during the study, we had each participant's initially set to the Default mode for the first week. This mode replicated the

Table 1. Assignment of each group and order

Condition	Description	Number of participants
Group Manual order 1	Default ->Manual ->Hybrid	7
Group Manual order 2	Default ->Hybrid ->Manual	8
Group Auto order 1	Default ->Auto ->Hybrid	7
Group Auto order 2	Default ->Hybrid ->Auto	8

original notification order, which placed IM and email notifications toward the top of the order, according to an official Android Developers article regarding people notifications[4].

During the three-week study period, participants were divided into either the manual group or the auto group. The two groups utilized both the Default mode and the Hybrid mode for a week each, with the difference being whether the other week was in the Manual mode or the Auto mode. Irrespective of their assignment to the manual or auto group, every participant engaged with all functionalities. The purpose of this grouping was to facilitate a thorough comprehension of users' preferences and perceptions regarding manual and automatic functionalities.

Following the Default mode, in the remaining two weeks of the study, each participant experienced two additional modes, one in each week: the Hybrid mode and either the Auto mode or the Manual mode. In order to reduce order effects, we counterbalanced the conditions and randomly determined the sequence of modes for each participant. Table 1 shows the number of participants in each group and order.

When using the Auto mode, NotiManager provides participants with the functionality of automatic sorting and automatic categorization of notifications. Conversely, in the Manual mode, participants have manual capabilities such as manual sorting of notifications, pinning, manually adding a notification to a category, and manually adding new categories. In the Hybrid mode, all functionalities from both the Auto mode and the Manual mode are integrated.

4.2 Experience Sampling and Diary Study

The objective of the ESM study [69] was to gather participants' in-situ usage, perceptions, and experiences with the features, as well as their reasons for using these features in various contexts. The diary study was aimed to collect more reflections from participants about their overall perceptions and interactions with these features throughout the day.

4.2.1 Research Instrument. NotiManager not only served as an app for users to view and handle notifications, it also recorded notifications that arrived on users' smartphones, tracked the usage of notification management features in NotiManager, determined whether to trigger ESM questionnaires based on user interactions with NotiManager, triggered diary questionnaires at a specific time, and logged phone-sensor data.

4.2.2 ESM Questionnaire. Each page of the ESM questionnaire was organized into several sections (see Figure 2a). The questions varied across different modes, catering to the unique features of each. In the Default mode, designed primarily to acquaint participants with the process of viewing and interacting with notifications in NotiManager, questionnaires were still utilized to familiarize participants with the process of completing them. In this mode, participants were asked to rate the extent to which the notification order at a specific time matched their ideal sequence. They also evaluated how well the system ordered their notifications in terms of urgency and importance, using a five-point Likert scale for all rating items. The questionnaire also gathered context-specific information about when participants encountered these notification sequences, including location [23, 76], activity

[4]The official Android Developers publication on Medium – People notifications: https://medium.com/androiddevelopers/people-notifications-2a2e4fb6ee96

[19, 22, 48, 50], their level of engagement in that activity [40], whether they were interacting with others at the time [9, 54, 64], and their degree of busyness during that moment [79].

In the Auto mode, participants were presented with comparative questions regarding two notification orders to evaluate the outcome from the automatic sorting feature (see Figure 2b). Specifically, they were asked to choose the order that most closely matched their ideal order. An additional response option, "The two orders are similar," was also provided. The first order replicated the arrangement in the current notification system, typically placing IM and email notifications towards the top [4]. The second order was generated based on predictions from our model. To minimize potential bias, the sequence in which these two notification orders were presented was randomized, ensuring participants were unaware of which order was the original and which was adjusted by our model.

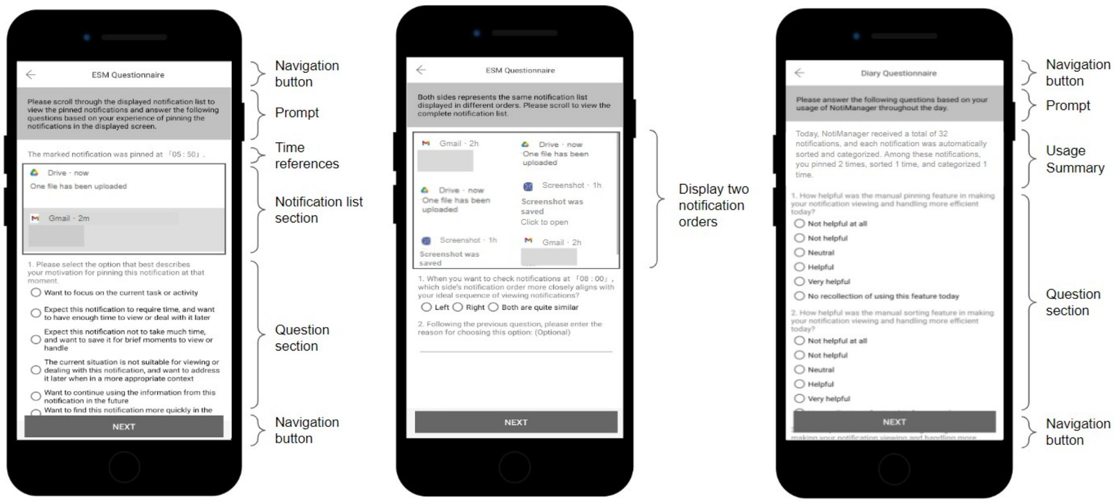
In the Manual mode, if participants had not utilized any of the three features since the last ESM questionnaire, they were asked to rate the extent to which the notification order at a specific time point matched their ideal sequence. If they had used any of the features, they reported their reasons for using them, the perceived necessity of using these features, and to provide an assessment of the notification that utilized the features. This assessment included evaluations based on the importance, urgency, and the context in which they used the features.

In the Hybrid mode, the questionnaire included question types from both the Auto and the Manual mode. To maintain focus, the questionnaire was limited to reporting on a maximum of three features. When selecting which features to include, diversity was prioritized to ensure that each type of utilized feature was represented at least once.

When querying participants about their reasons for using the features, we instructed them to select only one option that best described their reason. The choices were derived from Chang et al.'s study [11], reflecting users' intentions for checking notifications in specific activity contexts, and included: 1) To focus on the current task or activity; 2) To defer notification to convenient time; 3) To defer notification to spare time; 4) To defer notification to post-context; and 5) To locate the notification more quickly in the future. We included additional reasons for the categorizing feature: "To categorize notifications for later processing", and for pinning function: 1) To retain information for future reference; 2) To view the content of the notification without having it deleted; and 3) To prevent the notification from being deleted. These reasons were identified through our pilot study and personal use of NotiManager.

4.2.3 Diary Questionnaire. The diary questionnaire was structured with several components on each page (see Figure 2c). It employed a five-point scale to assess the perceived helpfulness of each feature. Following the questions on the helpfulness of these features, a text-box question allowed participants to provide detailed feedback on their overall feelings and experiences of using NotiManager throughout the day.

4.2.4 ESM and Diary Mechanism. Upon installation of NotiManager, participants selected a time-window of at least 12 hours during which they were willing to receive ESM and diary questionnaires each day. NotiManager triggered ESM questionnaires based on specific criteria: 1) the sampled time within the user-defined time window; 2) at least one hour had passed since the last questionnaire completion, or at least 30 minutes had passed since the last questionnaire was prompted; 3) any use of management features, or if none are used, changes in the displayed notification order occurred within the last 30 minutes, a time threshold derived from prior research (e.g., [12]). Meeting these conditions, NotiManager promptly dispatched ESM notifications to participants' native notification drawers when they closed NotiManager, ensuring that these notifications didn't blend with other notifications in NotiManager. The diary questionnaire was scheduled to be sent every night, approximately one hour prior to the user-defined time window. Given the complexity of the ESM and diary questionnaire and the mechanism, we conducted a pilot test with three participants to validate the questionnaire and its functioning. The pilot study streamlined the formal study process, lasting six days with each mode being utilized for two days.



(a) (b) (c)

Figure 2. The ESM questionnaire interface (a), an example of ESM about comparative questions regarding two notification orders (b) and the diary questionnaire interface (c)

All three pilot participants could complete the ESM and diary questionnaire within an average of three minutes, an acceptable duration for an ESM study [7, 18, 20], we then proceeded with the formal study.

4.3 Study Procedure

Each participant attended an in-person pre-study meeting during which we assisted them in installing NotiManager on their phones, introduced the interfaces, and walked them through the study procedure. We did not ask participants to use any specific features; instead, we informed them about the availability and characteristics of each feature, leaving them the freedom to decide how and when to use them as they saw fit. Our aim was to observe whether and how these features would naturally integrate into or augment their existing notification management practices. Over the three-week study period, participants experienced a different mode of NotiManager each week. Modes were automatically switched after a week, with participants being notified of the impending change via email and an in-app pop-up window that confirmed the update. At the conclusion of the study, we invited participants to participate in semi-structured interviews to deepen our understanding of their experiences. Of those invited, 28 participated in these interviews. During the interviews, we focused on exploring their usage, experiences, and strategies with NotiManager. We particularly probed the circumstances under which participants adopted specific strategies or held certain expectations, aiming to understand the underlying needs and reasons. We also solicited feedback on potential improvements and their ideal vision for NotiManager, particularly when they expressed dissatisfaction with current features, with an aim to uncover the needs behind their expectations and identify areas for enhancement in future versions of the system.

Compensation was based on the number of completed ESM and diary questionnaires, at US$0.3 each. Participants' usage of features did not affect their compensation. The duration of participation was also considered. Full participation in both the first and second weeks earned US$12, and completing the third week added US$6. Additionally, those who took part in the post-study interview received an extra US$6. The study received approval from our university's Institutional Review Board (IRB).

4.4 Recruitment and Participants

We primarily recruited participants through various Facebook groups dedicated to discussions and usage of different brands of mobile phones, as well as by connecting with potential participants through local research networks in our country. Each recruitment ad included a link to a sign-up form with questions covering subjects such as the average daily count of notifications, the top five notification categories they receive, and their behaviors in relation to the notification drawer. This approach was chosen to ensure a diverse range of participant backgrounds. We selected our participants who 1) checked their notification drawers and handled notifications every day; 2) received an average of more than 10 notifications per day; and 3) used their phones for over 2 hours daily. This resulted in 30 participants, all of whom participated in the study for a full 21 days. The participants' ages ranged from 18 and 40 (M = 26). This diverse group included 21 students and 9 non-students, with a gender distribution of 18 females and 12 males.

4.5 Quantitative Data Cleaning and Analysis

We recorded 109,253 notifications in total and received 2,090 ESM responses along with 506 diary responses, with a response rate of 71% and 80%, respectively. We excluded 36,756 notifications that emerged when users were using the Default mode, as the initial week was primarily intended for participants to become accustomed to using NotiManager. This left us with a dataset comprising 72,497 notifications from 424 unique apps (for further details on the types of these notifications, please refer to the appendix A). The majority of these notifications were dispatched by three leading messaging apps: *LINE* (21,336 notifications; 90% of participants), *Telegram* (8,566 notifications; 20% of participants), and *FB Messenger* (6,823 notifications; 87% of participants). We also discarded 672 ESM responses and 178 diary responses that were filled out in the Default mode. 38 diary responses indicated not remembering using any feature on this day were also removed, leaving 1,418 ESM responses and 290 diary responses available for further analysis. 37% of pinning events, 40% of manual sorting events, and 24% of manual categorization events were asked and responded to in the remaining ESM responses. Using the remaining notification data and responses to explore user preferences between the automatic and manual features and to gauge the perceived helpfulness of individual features, we conducted a logistic regression analysis with the "lmerTest" [36] package in R software[5], and included participant ID numbers as a random effect to account for individual differences among the participants.

4.6 Qualitative Data Analysis

We transcribed the interview recordings and conducted data analysis simultaneously with the data-collection process using Atlas.ti[6], an online collaborative qualitative-analysis software. We employed an inductive thematic analysis approach [6], wherein the development of the codebook was an iterative process. Two members of our research team took on the role of coders and started by open coding to explore important concepts in the first three interview transcripts. Subsequently, the coders participated in multiple collaborative sessions during which they discussed, compared, and merged their individual codebooks to reach a consensus on the initial set of codes. This process aimed to enhance consistency and mitigate individual biases [46]. As data collection progressed, the researchers used these preliminary codes to independently code the transcripts. Throughout the coding process, the researchers consistently compared data, codes, and memos. They incorporated their epistemological beliefs in the analysis process, including their own reflection and interpretation of the data. They engaged in discussions to refine and revise codes iteratively as necessary. Whenever new codes were introduced, modified, or refined, their applicability was evaluated against other data until consensus was reached. This process resulted

[5]R software: https://www.R-project.org/
[6]Atlas.ti: https://atlasti.com/

in the emergence of numerous themes that pertain to how different automatic and manual features influenced participants' behaviors and perceptions in notification management.

5 QUANTITATIVE RESULTS

On average, each participant accessed NotiManager 27 times per day (Md = 21; SD = 20.91). Pinning was utilized the most (230) among the manual features, followed by manual sorting (162), and manual categorization (33). Regarding perceived helpfulness, which is rated on a five-point scale, we observed a similar order: pinning was also rated as the most helpful (M = 3.71, SD = 0.81), followed by manual sorting (M = 3.54, SD = 0.82), and manual categorization (M = 3.41, SD = 0.88). We provide a detailed analysis of each feature's usage below.

5.1 Pinning Notifications

We recorded a total of 230 pinning events involving 219 unique notifications from 27 different apps. Out of the 30 participants, 22 (73%) engaged in the act of pinning notifications. In the Manual mode, participants used pinning an average of 1.26 times per day, compared to 1 times per day in the Hybrid mode. Notifications related to communication dominated the pinning activity, with 57% originating from IM applications.

There were diverse reasons for pinning notifications. As depicted in Figure 3, the primary reason, accounting for 31%, was to defer addressing the notification to a convenient time. The second most cited reason was to prevent accidental deletion of the notification (27%). The third most cited reason was to facilitate easier access to the information in the notification later (15%). The diversity in reasons for pinning notifications was mirrored in the variable durations for which notifications were pinned, as illustrated in Figure 4 and Figure 5. When notifications were pinned for the reasons for deferring to a convenient time, the average pinned duration was 716.31 minutes (SD=978.67). However, when pinning was for protecting notifications from accidental deletion, the average duration was notably shorter, at 273.35 minutes (SD = 586.52). The longest duration of pinning, averaging 1,690.26 minutes (SD = 1634.51), occurred when the reason was to facilitate future access to the information.

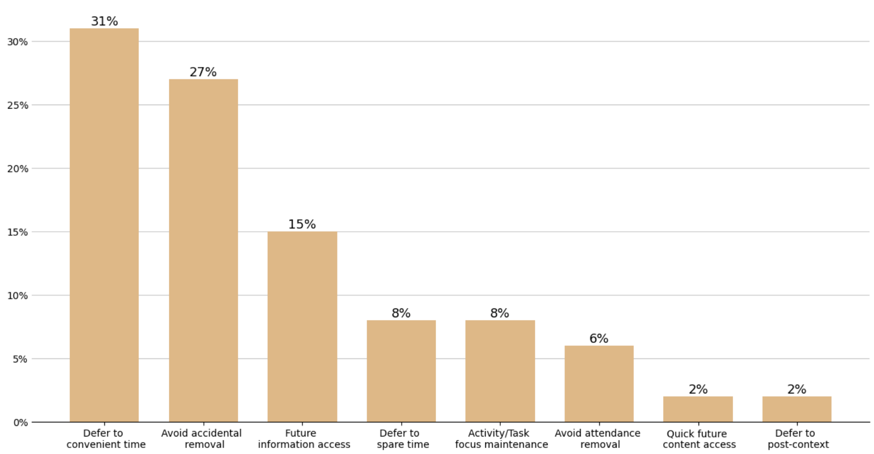


Figure 3. Participants' self-reported motivations for pinning notifications. (Only one motivation was provided for each pinning instance.)

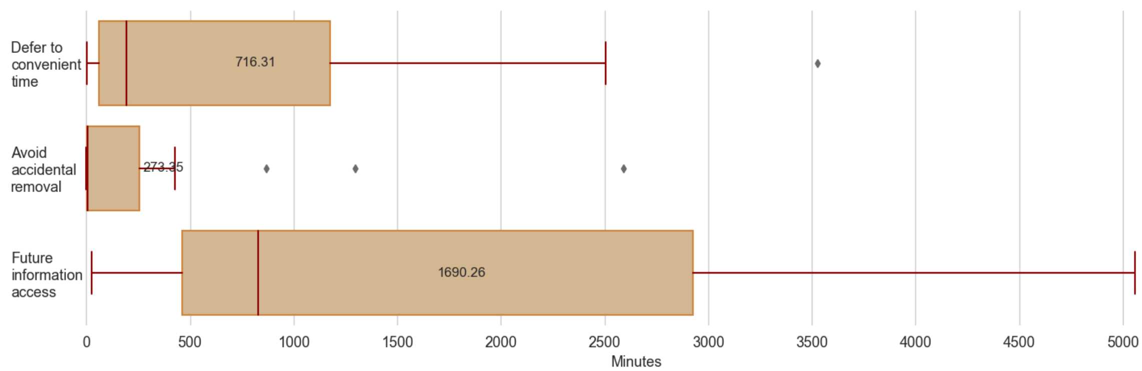


Figure 4. The duration of pinning corresponds to the top three selected reasons for pinning. (Defer to convenient time: Q1=60.45, Q2=194.21, Q3= 1176.08; Avoid accidental removal: Q1=2.51, Q2=8.06, Q3=255.04; Future information access: Q1=464.22, Q2=829, Q3=2927.93)

Logistic regression analysis further substantiated that the duration of pinning for this reason was statistically longer than when the pinning was aimed at protecting the notification from accidental deletion ($t(65.96) = 2.188$, $p = 0.0322$, SEM = 367.4).

In addition, pinning may have been utilized for bulk-removal on other notifications. We observed that when notifications were pinned to prevent accidental deletion, more than four-fifths (83%) of these pinned events were followed by a click on the clear-all-notifications button before exiting NotiManager, which would remove all not-pinned notifications. When the reasons for pinning were to facilitate future access to information or defer notifications to a convenient time, the actions of deleting all notifications followed by pinning were 7.7% and 64% respectively. The subsequent action of deleting all notifications when pinning to prevent accidental deletion was in stark contrast to when the reason was to ease future access to information (Z = -2.337, p = 0.0194, SEM = 1.6301).

5.2 Sorting Notifications

A total of 162 manual sorting events were recorded from 140 distinct notifications spanning 40 different apps. Among the 30 participants, 22 (73%) of them made use of the manual sorting feature. Under the Manual mode, participants utilized manual sorting an average of 1.93 times per day, whereas under the Hybrid mode, the average was 0.52 times per day. Participants most often adjusted the position of communication-related notifications. The IM category composed 46% (64) of all manual sorting notifications. Notably, nearly one-fifth (19%) of the manual sorting actions were on pinned notifications.

Figure 6 shows the primary reasons participants manually adjusted the position of notifications. The most prevalent reason, accounting for 38% of instances, was to facilitate easier location of the notification later. The next most common reason, making up 19% of cases, was the expectation that addressing the notification would not require a significant time commitment, leading participants to defer it to spare time. This was followed by the desire to maintain focus on the current task or activity, cited in 14% of cases, also relating to deferral. The direction in which participants sorted notifications varied by their underlying reason. For quicker future access, all manually sorted notifications were moved upward. Conversely, when deferring notifications to a spare time, 83% were moved downward.

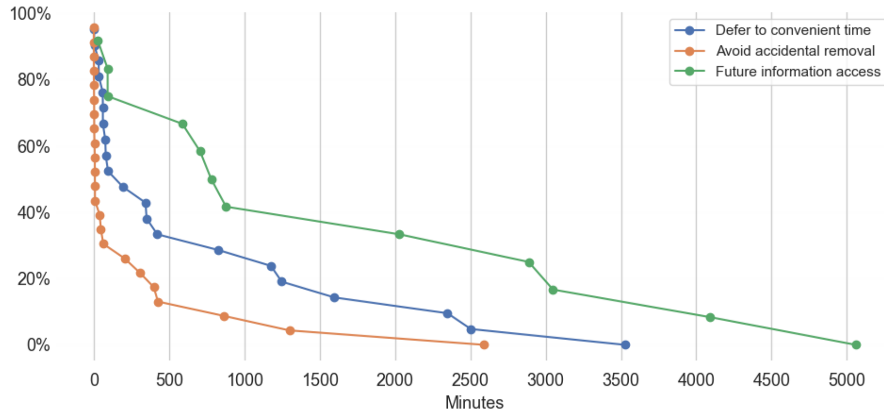


Figure 5. The actual distributions of pinning duration corresponding to the top three selected pinning reasons. (5 instances of deferring to convenient time, and 1 instance of future information access were removed as they were remained pinned until the end of the study)

In evaluating the automatic sorting outcomes from 864 instances, the original notification order was preferred over the automatic sorting 62.5% of the time, while the automatically sorted list was chosen only 13.77% of the time; the remaining 23.73% found both orders similar. Furthermore, ratings from participants indicated that manually sorted notifications were closer to their ideal order (M = 3.26; SD = 0.82) compared to those sorted automatically (M = 3.01; SD = 0.87, $t(493.6832) = -2.337$, $p = 0.0198$, SEM = 0.1689).

5.3 Categorizing Notifications

Among all participants, only one (P29) manually created a new category in the Hybrid mode, using it just twice for categorizing notifications and accessing it three times. In contrast, a larger number of participants utilized the system's built-in categories, derived from automatic categorization, to classify and view notifications. Specifically, in the Hybrid mode, 7 participants (23%) manually categorized notifications using these built-in categories an average of 0.63 times per day. 24 participants (80%) accessed notifications via automated categories. In the Auto mode, they did so at an average rate of 1.66 times per day, while in the Hybrid mode, the average was 1.04 times per day. The most clicked category was *IM* (22%), followed by *Social* (12%) and *Transportation* (11%). Given the limited usage, we only received limited number of self-reported responses for using categorization (n=8). Half of these responses were to collectively process notifications that belong to the same category.

Gaining the overview of how often and why the participants used the three features in the next section we reported participants sharing of how to utilize these features in your daily lives during the interviews these include how they perceive these features experience with them strategies of using them and why they found certain features as not useful as expected which allow us to unconvert the needs underlying this usage and expectations. In the following sections we divide our findings by features.

6 QUALITATIVE FINDINGS

Having provided an overview of the usage and reasons why participants used the three features based on app logs and ESM responses, this section delves deeper into participants' perceptions, experiences, strategies, and expectations regarding these features. Our findings are organized by feature in the subsequent sections.

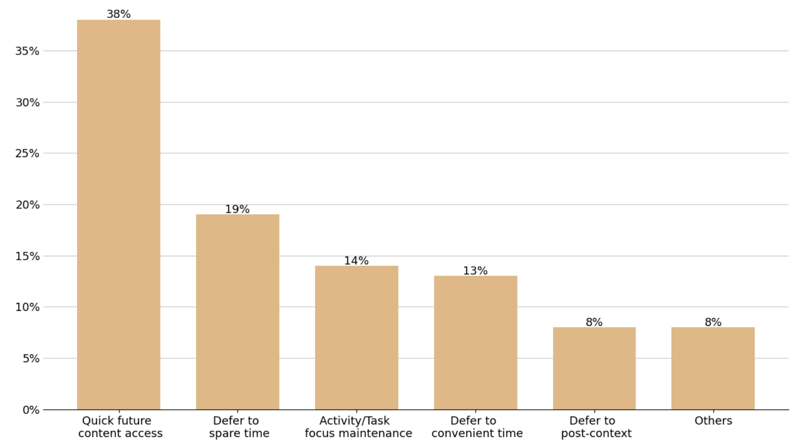


Figure 6. Participant's self-reported reasons for sorting notifications

6.1 Categorization: Automatic Over Manual and Beyond App-Based Methods

Overall, participants found automatic categorization useful but felt it should extend beyond mere app-based categorization. Manual categorization was seldom used and, when employed, was primarily intended to enhance the automatic feature rather than for organizing notifications themselves.

6.1.1 Automatic App-Based Categorization Facilitates Notification Operation, but Topic-Based Mechanism is More Adequate. Many participants expressed a preference for processing similar or related notifications collectively, such as reading or deleting a specific type of notifications, as it allows for more efficient operation. Therefore, some participants utilized the built-in categories of the system for this reason. For instance, P14 noted, *"Sometimes I found it useful to clear a specific type of notification altogether [by using the system built-in categories]. For instance, shopping websites often send numerous messages, so I opened the shopping category and deleted them together."*

However, other participants found app-based categorization employed by NotiManager inadequate and expressed a desire for a topic-based approach. They pointed out that notifications from a single app could encompass a variety of topics, as P4 explained, *"If Instagram sent notifications about someone liking my post, it seemed appropriate to place these notifications under social media. However, if it sent notifications about someone responding to my story, I might question why it was not categorized under instant messaging."*

Furthermore, they found notifications concerning the same topic were also frequently dispersed across multiple categories, complicating batch operations based solely on the app. P27 indicated, *"I felt that important things may come through email, Line, or SMS. These three apps were not in the same category, so if I wanted to check the notifications sent from these important apps, I had to check three categories."* To improve the utility of automatic categorization, many participants advocated for a topic-based mechanism that groups notifications by topic rather than by the originating application. This approach would enhance the organization and management of similar or related notifications, making it easier for users to handle their notifications more efficiently.

6.1.2 Manual Categorization was Tedious. Used Mainly for Correcting Automatic Feature. Participants rarely created new categories or manually moved notifications to specific categories. An exceptional case involved P29, who initially experimented with adding new categories but ultimately found the process overly complex and lacking in benefits. Similarly, P23 expressed reluctance to manually categorize notifications due to the tediousness of the process, explaining, *"There were a lot of messages at that time, and I found it quite difficult to categorize each one individually."* In the Hybrid mode, where automatic categorization was available, participants felt even less incentive to categorize notifications manually. As P6 stated, *"It [automatic categorization] already categorized them very thoroughly, I rarely found the need to do manual sorting myself. It was almost unnecessary."*

Notably, some participants indicated that their changes to categories were primarily made to correct the outcomes of automatic categorization, rather than for organizing notifications. For instance, P14 noted a specific instance of re-categorization, stating, *"I had a frequently used IM app called Mazemoster. Perhaps because it was not very common, it didn't get categorized under the IM category but was instead placed in the Other category. I thought changing its category to IM might help the system prioritize its notifications in future sorts."* Notably, this motivation stemmed from the assumption that the research app's categorization mechanism would learn from user input, similar to many social media platforms. Once it was clarified that this adaptive learning was not part of the system, participants expressed a desire for future systems to allow users the ability to adjust the system's behavior or overwrite its categories. Despite recognizing that this would require extra effort, they were willing to undertake it to ensure the categorization outcomes better aligned with their desired results.

6.2 Pinning: Highly Appreciated and Utilized for Various Purposes

Our interview results resonate with the common use of the pinning feature in the quantitative results, with many participants mentioning their appreciation for it and its various applications. Additionally, they often compared it to the snoozing feature, noting that its function extends beyond mere deferral.

6.2.1 Pinning was Useful for Deferral, Reminder, Handling Large Volume of Notifications, and Reserving Information. Participants provided various reasons for using the pinning feature during the interviews, some of which revealed nuanced ways of utilizing pinning to facilitate task and time management.

(1) **Deferral while Maintaining Reminder:** The most commonly cited reason for using the pinning feature was to defer handling notifications. Common reasons included notifications requiring substantial time or effort, the situation being inappropriate for immediate handling, or the preference to read or respond on a more suitable device. For instance, P20 explained,*"I noticed that the notification contained multiple messages. So I decided to respond later when I had more time to focus on it."* Similarly, P18 used pinning to delay downloading files until he could use his computer, stating,*"It was a music service sending me an email asking me to download music files, and so I pinned it. [...][When I was sitting in front of my computer], I opened Gmail on my computer to download these things."*
    Additionally, many participants highlighted that pinning not only served for deferral but also serves as a visual cue, consistently reminding them of deferred tasks each time they accessed NotiManager. As P21 indicated, *"The notification told me what I needed to do today, and I had already seen it. At that moment, I also had some free time, but I might just be afraid of forgetting, so I pinned it there. This way, every time I opened the notification drawer, it reminded me again."* Participants also mentioned that the presence of an icon next to pinned notifications further enhanced the perceived importance of the pinned notifications, thereby priming them to pay special attention.

(2) **Facilitation of Notification Handling Workflow:** Participants found the pinning feature particularly useful when dealing with a large volume of notifications. This feature allowed them to temporarily preserve notifications that seemed "potentially" worth more attention while quickly scanning through. They could then use the clear-all-notifications button to remove the rest, and later return to these pinned notifications to review their content in detail and decide whether to leave them pinned or to unpin them. For example, P22 described this strategy, stating, *"The main idea is to first keep the notifications that need to be dealt with and clear all the others from the notification list. This visually makes it more comfortable. Then, if there's something that needs to be handled immediately, I would unpin that notification and address it."*

Many participants appreciated this workflow and mentioned that it has become their standard method for managing notifications, because it enabled them to quickly narrow down to specific notifications, minimizing interface clutter by batch removal without the risk of accidentally eliminating important ones. Furthermore, several participants particularly valued this workflow because, in the past, they had to leave notifications potentially worth their attention *unread*, as clicking into them would remove them from the notification list. Additionally, these unread notifications were not visually distinct from new ones, and keeping them in the list prevented the use of the clear-all function, leading to a mix of new and old notifications over time. The availability of the pinning feature alleviated the need to accumulate unread notifications in the list, allowed for the use of the clear-all feature without concern, and enabled them to revisit the same notification multiple times. As P2 noted, *"Sometimes, deleting messages one by one can be too slow, leading to the temptation to use the clear-all-notifications option. However, this could sometimes result in accidentally removing important messages. I think pinning really helps prevent these messages from being removed."*

(3) **Information Centralization:** Participants also perceived pinning as a convenient method for reserving information contained in notifications. By using the pinning feature, they could centralize their information management within one application, NotiManager, thus reducing the effort needed to preserve and locate information across various platforms. For instance, P15 noted,*"In the past, I had to open the browser first when I wanted to revisit a specific webpage. Then, I located the URL by using the browser tab. With the pinning feature, I don't have to go through these steps. I can just pin the notification and when I have time to check the webpage, I simply click on the notification."* Participants highlighted that quick access was convenient for constantly updated notifications, such as group chats or regular updates, due to the frequent need to access these threads. Pinning these notifications created a convenient portal for revisiting them. P6 described,*"I could go in and handle it at any time. I wanted to keep this notification as a chat window here [in the app], and on the top."* Finally, participants valued pinning for its ability to preserve complete notification details, helping them remember the original reasons for saving that information, especially if they intended to address these notifications after a long period.

6.2.2 Pinning was Considered Poor for Proactive Reminder, but Less Disturbing and More Flexible than Snoozing. Reflecting on the differences between pinning and snoozing features, participants highlighted several key distinctions. The primary difference noted was that snoozing acts as a system-initiated reminder, automatically alerting the user after a set period, whereas pinning relies on user initiative to remember and revisit notifications. therefore the former allows users to be passively reminded by the snooze feature, freeing them from the need to proactively check the NotiManager app. This feature was appreciated for providing proactive reminders when needed. For instance, P20 stated, *"Sometimes I completely forgot about my pinned notifications. It was only when I revisited the notification app that I rediscovered the task. But if I had set a specific snooze period, the notification would trigger the alert again by itself, which would be a good reminder."* However, some participants felt that the alerts from snoozing created additional pressure, as P29 described, feeling *"like I'm being chased by someone."* Others mentioned that these alerts could be confusing, making it unclear if the alert was from a new notification or a snoozed one, or simply added disturbance, as P19 commented, *"If it alerts a second time, and I only discover it's the same notification as before, I would find it disturbing."*

Moreover, several participants noted challenges in setting a suitable snooze duration, often needing additional snoozes if they weren't available within the initially set period. In contrast, pinning was valued for offering greater flexibility in managing notifications, particularly when the exact time to address them was uncertain. P10 illustrated this flexibility, stating, *"With pinning, if I initially anticipate being free two hours later but find myself available after just one hour, I can handle it directly. However, with snoozing, I must wait the full two hours before addressing it."* Thus, while pinning generally offered better flexibility, snoozing was favored when participants required explicit and intrusive reminders to ensure they would not forget to address notifications within a specific timeframe.

6.3 Sorting: Advocating Enhanced Automatic while Appreciating Manual for Adjustment and Feedback

Overall, the general sentiment among the participants was that the automatic sorting outcome did not align well with their preferred order most of the time, and that they found availability of manual sorting beneficial. More details are presented below.

6.3.1 Automatic Sorting was Expected to be Supporting Batch Operation, Adaptive and Explainable. Our qualitative analysis identified three main issues that participants perceived with the automatic sorting feature of NotiManager, highlighting areas for improvement.

First, participants noted that the current automated sorting of NotiManager did not facilitate efficient batch operations on notifications, a limitation they found also true for existing notification systems. Specifically, they expressed a desire for the system to sort notifications based on the similarity of topics or themes. This would allow them to maintain a consistent mindset while handling batches of notifications, thereby reducing the need for cognitive switching. As explained by P5, *"When I saw a work-related notification, I approached it with a work mindset, and if the next notification was unrelated to work, I would switch to a relaxation mindset. If notifications were sorted by type, I could handle all work-related notifications at once before shifting to more leisurely ones."* Similarly, P6 also noted, *"I saw a block of notifications from LINE here, another block of notifications from LINE there, then some unimportant notifications from Foodpanda in between, [...], not really convenient if you got to check so many blocks."* Additionally, participants mentioned that a mix of topics in notifications could lead to important messages being overlooked, especially if they were buried among less critical notifications. For instance, P20 shared her experience with and concern about missing crucial notifications, *"I could easily overlook important notifications, like those from teachers or classmates, when they were mixed among a bunch of ads. I might skip this entire batch of notifications."*

Second, participants deemed the current automatic sorting of NotiManager limited in its adaptability, particularly wanting a sorting function that could adjust according to their activity context and individual preferences. Discussing adaptability to context, P10 stated, *"Most of the time, I might reply to this work-related group chat immediately. But when I'm off work or during weekends, I don't need it to be prioritized at the top of the notification app."* P23 added, *"There are some days I just don't want to do work in the lab. On those days, I would prefer those notifications to be at the bottom of the app. But the system probably considered them important and put them first."*

Participants also expressed a desire for configurations to express their preference. For instance, P27 explained his preference for setting this during onboarding, *"Maybe the system could ask what kinds of notifications I preferred when I installed this application for preliminary personalization. Most apps require configuration right after installation, so I would be more patient to do this than when I'm actually using this app."* Several participants also expected the system to learn their input could impact the system to improve its sorting performance. For example, P6 shared her desire to manually sort notifications to express her personal preferences for future enhancements. P27 anticipated that the system could learn from the order in which he attended to notifications, noting, *"It could improve by the order I click notifications, just like how keyboard input selection gets optimized."*

Third, several participants expressed a desire for a clearer understanding of how NotiManager sorted notifications, as they struggled to make sense of the underlying logic. For example, P21 illustrated this confusion by stating, *"Sometimes it appeared to be sorted by the importance of the app associated with the notification, while other times it seemed to be sorted by time. It's a bit messy."* With a better knowledge of the system's operation, they believe they could, as P18 suggested, *"have a rough picture of where a notification was, if provided with some description,"* or adjust the settings to better align with their preferences.

6.3.2 Manual Sorting was Used for Control Visibility, Grouping, and Correcting Automatic Feature. Our interview revealed three main uses of manual sorting by the participants. First, participants used manual sorting to control the visibility of pinned notifications. A common practice was placing to-do items at the top of NotiManager's notification interface to prioritize tasks. P7 described this practice, stating, *"I would pick out which notifications I should deal with on the coming day every time I went to bed. [...] On the following night, I would dismiss those I had completed and drag upwards the ones for the next day."* Although the pin icon already added salience to the notifications, placing them at the top further enhanced their visibility, as P18 noted, *"It's like the starred emails that go on top in Gmail. [...] If they're important, I want to have more chances of seeing them."* Conversely, for notifications they intended to handle later but did not want obstructing the visibility of new notifications, participants positioned these at the bottom. This approach ensured that they *"must have reviewed all new notifications above them,"* ensuring nothing was overlooked.

Second, participants sometimes used manual sorting to organize information for easier future access. P23 mentioned clustering important notifications together and separating this cluster from less important ones, aiming to *"be able to see important notifications grouped together."* She further suggested, *"I think there should be a clear border between important notifications and those that aren't so important."* Third, as mentioned earlier, several participants used manual sorting to provide feedback to the automatic sorting feature, assuming it would learn from their behavior.

7 DISCUSSION

In previous sections, we have detailed participants' usage (and non-usage) of the three features, along with their perceptions, experience, strategies, and expectations around them. In this section, we discuss key takeaways from these results and their implications for future notification systems.

7.1 Pinning, as a Simple Feature, Reveals Various Needs in Notification Interaction

In both our qualitative and quantitative results, participants reported various reasons for using the pinning feature on their notifications during the study. While pinning is a simple action—achieved by clicking on the pinning icon on a notification—it reveals multiple vital needs that future notification systems should support. Specifically, we first identified that deferring the reading and response to notifications was one of the main reasons participants used the pinning function, confirming Lin's assumption [43] that this feature can help address these needs. In particular, our findings also clarify how users might perceive and use pinning and snoozing differently for deferral in their daily lives. One of our observation aligned with Weber et al. [73]: namely, that participants might not always be able to estimate an accurate duration for deferral. However, our findings add further insight into the perceived pros and cons of each feature. For example, participants generally perceived that alerts prompted by the re-emergence of snoozed notifications functioned as proactive reminders. However, while this was useful when a prominent alert was needed, at other times participants felt it created pressure, as though they were being "chased" by their tasks. In contrast, pinned notifications were seen as less intrusive, providing a constant visual reminder in the notification interface. Participants especially appreciated the flexibility of pinning that allowed them to decide when to address different notifications. And because it does not issue additional alerts, participants deemed that they would not be confused with incoming new notifications.

Crucially, participants leveraged the pinning feature not merely for deferral but also for other reasons. For example, some used it to ensure easy and constant access to information, appreciating its advantage of centralizing information needed for future use in one single place. Additionally, participants appreciated its ability in preserving all information related to the pinned notifications that helped them contextualize what to do with the notifications.

What we found even worth noting is the usage of pinning for preventing accidental deletion of important notifications, as reported in the ESM, which we revealed from the interview result, was primarily driven by the desire to quickly clear the interface of notifications deemed unimportant, thereby allowing focus on those requiring more thorough assessment. Specifically, to perform this efficiently, participants preferred using the one-click-delete-all button — a method also noted in [74] — rather than swiping off notifications one by one. Yet, this bulk removal method risks the accidental deletion of important notifications. The pinning feature, by enabling users to temporarily preserve those deemed potentially important, allows users to quickly narrow down to specific notifications by performing a bulk removal of all others, without the need to leave notifications unread and worry about accidental deletion.

This practice, i.e., temporarily preserving notifications before bulk removal mainly for later assessment, is evident in their brief pinning duration compared to other uses. Experiencing this advantage, many participants reportedly had established a new notification management workflow around this usage. We found the emergence of this workflow particularly interesting and noteworthy, as it reflects, to some extent, the limitations of current notification systems in supporting the need for efficiently narrowing down notifications while alleviating concerns about accidental deletion.

Additionally, in current notification systems, interacting with a notification causes it to disappear from the interface. To avoid this, participants often left notifications unread in the notification list when they wanted to defer reading them. However, these unread notifications are visually indistinguishable from new ones, creating confusion and posing a risk of accidental deletion. We observed that participants' appreciation of the pinning function also largely stemmed from its ability to eliminate the need to leave notifications unread, thus simplifying the process of distinguishing between deferred notifications and new ones.

In summary, the usage of the pinning feature reveals several vital needs in how mobile users often want to handle their notifications. These include deferral without additional alert or mixture with new notifications, having easy and constant access to information for later use, quickly narrowing down notifications worth attention on a clean interface without worrying accidental removal, and adding visual salience to help distinguish notifications. To support these needs, we suggest that future notification systems include a pinning feature, and meanwhile, devise ways to effectively support each of these needs. While pinning is a useful and simple approach, it may not be the optimal solution for addressing all these needs comprehensively. Additionally, there are occasional needs such as proactive alerts for reminders that pinning fails to address but are covered by other features such as snoozing [73]. This suggests that future systems could consider incorporating pinning with other features, allowing users to apply them based on their specific intentions.

7.2 Towards Intelligence: Implications for Future Automated Sorting and Categorization

Drawing from prior research, we hypothesized that automated sorting and categorization could enhance users' efficiency in managing notifications by prioritizing and filtering those of greatest interest, as suggested in [43]. Although the original intention behind implementing the automatic sorting and categorization features was not to create an intelligent system—thus being simplistic in nature—but to prompt participants to discuss potential improvements and express expectations, it was still somewhat unexpected that automatic sorting received a lower assessment than the default sorting. Nevertheless, thanks to the flaws participants perceived from NotiManager's automatic sorting, we successfully extracted users' needs and desires regarding how systems should order and categorize notifications. We discuss these results and their implications in further detail below.

7.2.1 Grouping Notifications with Similar Topics is Useful for Both Sorting and Categorization. Our findings suggest that participants highly valued being able to efficiently perform batch processing and operation on similar or related notifications, as it reduces the need for repeatedly performing similar actions and reduces cognitive switching. Currently, both NotiManager and existing notification systems fail to group topics and semantics embedded in notifications, relying mainly on app information. Therefore, we recommend that future automatic sorting and categorization for notifications prioritize semantic content over merely app-based information. While previous work has developed novel machine-learning approaches for content-driven intelligent notification systems [41, 47, 49], the advent of machine learning models on semantic analysis [8, 53, 65] presents a new opportunity to significantly improve this area. For example, large language models (LLMs) have shown great promise in analyzing semantics [33, 72, 78], and utilizing them could enable more accurate categorization of notifications with similar topics. However, given that the efficacy of these models in analyzing notification similarities based on topic remains untested and could prove challenging, provided that perceptions of topics can vary among users, We suggest that future research should investigate the performance of these models for this specific application and explore methods for fine-tuning them if they are found to be inadequate. In our study, due to the limitations of our categorization scheme, participants did not extensively use these categories to access notifications. Improvements in these categories may lead to more frequent use.

7.2.2 Adapating Notification Order to Context. Feedback on automatic sorting revealed participants' need to adapt the display order based on context, as they preferred to see different types of notifications in various situations. This finding aligns with recent studies suggesting that users prioritize different notifications in distinct activity contexts [11, 32, 43] and that intelligent notification systems with context sensing can more effectively predict suitable moments for delivery [47, 55, 56, 58]. Our empirical evaluation of automatic sorting based on machine learning further uncovered users' preference for dynamic sorting tailored to both topics and activity contexts. Therefore, our findings suggest that future notification systems should adapt notification sorting based on both types of information. Specifically, if the system employs LLMs or develops a specialized sorting model (e.g. [29]) to process notification information, it would be advantageous to incorporate contextual data, such as calendar events or activity information, alongside notification content to assess relevance and timing suitability for prioritization. However, if a general LLM encounters difficulties in prioritization due to the complexity of context-dependent preferences, future work could consider designing a human-in-the-loop mechanism to collect and incorporate user feedback across different contexts. This approach would enable the dynamic fine-tuning of LLMs for this specific prioritization task. Chang et al. [11] revealed users' intentions for engaging with notifications in specific activity contexts. We recommend that future research collect users' motivations for reading notifications using ESM, incorporating their intentions into the questionnaire. This could involve logging sensor data to train models for classifying reading intentions. These classifications could then inform the models, which would adjust notification sorting based on both the topic and the classified intentions. Subsequent research could investigate how well combining topic and intention classification meets user needs.

7.2.3 Improving Sorting and Categorization from Users' Input. Our study revealed that the use of manual sorting and categorization was often intended to correct automated processes, with participants anticipating that the system could recognize these corrections and adjust accordingly. This desire for the system to learn from their interactions indicates that participants were aware of the limitations of automatic sorting and categorization but were willing to instruct or make corrections to better align the system with their preferences and needs. Although NotiManager did not support this functionality, participants expressed a hope for future systems as they understood the challenges of fully adjusting to user needs initially, but believed the efficiency could be greatly improved if the system could learn from both their explicit inputs (corrections and instructions) and implicit inputs (natural interactions with notifications).

As a result, we suggest that future notification systems incorporate learning from both explicit and implicit user inputs. The explicit input could start with configurations upon onboarding, serving as a foundational step for the system to begin learning user preferences [5, 16, 32, 71]. Additionally, combining onboarding configurations with ongoing user feedback [16] could instruct the system to adjust to specific user behaviors in particular situations—a desire explicitly expressed by participants who utilized manual sorting and categorization intending to influence automated outcomes. Future systems could allow users to directly issue commands or provide examples to fine-tune the model to adapt to their desired behaviors. Alternatively, the system could prompt users for feedback or adjustments after interactions with notifications, implementing a human-in-the-loop approach to gradually refine system accuracy based on minimal user inputs. For implicit input, leveraging reinforced learning [15, 42, 80] from recommender systems could enable the system to learn from users' natural interactions with notifications, continuously improving notification sorting.

Lastly, as some participants expressed frustration at not understanding how the system works—which hinders their efficiency in locating specific notifications—we suggest that future systems, if leveraging automatic sorting or categorization, provide intelligible explanations within the interface. This transparency can help users understand the reasoning behind specific sorting or categorization decisions, thereby enhancing their ability to navigate the information more effectively [27, 34].

7.3 Study Limitations

The current study is subject to several limitations. First, because our notification management features could not be directly integrated into the Android notification drawer, participants were required to use NotiManager as the sole interface to access and manage their notifications, which could potentially influence how frequently they utilized these features. Second, even though participants had a week-long introductory period under the Default mode to become familiar with NotiManager, this period may not have been sufficient for them to fully grasp its functionalities. In addition, categorizing notifications involves a more complex process, possibly extending the time they need to get used to the feature. Third, the sorting model used in this study was based on a dataset from prior research [43], collected from a relatively small group of participants. The limited size of this dataset may impact the model's generalizability and effectiveness in accurately ranking notifications for users outside their participant pool. This limitation could contribute to the general dissatisfaction with the automatic sorting outcomes reported by our participants. Fourth, though the ESM triggering mechanism was refined through a pilot study, the eventual sample size of questionnaires for each function was relatively small. Fifth, although we provided detailed explanations during interviews, some participants might have lacked prior experience or a full understanding of how snoozing notifications function, potentially affecting their comparison with the pinning feature. Sixth, novelty effects might have influenced our feature usage statistics. Nevertheless, our quantitative analysis opted not to exclude usage data from the first days of the study since participants' daily usage statistics indicated no evident peak in initial usage nor a diminishing trend over time. Finally, our data were derived from a small group (n=30) of Android users within the authors' home country, with all study participants under 40 years old and more than two-thirds being students. Consequently, though our sample size and demographic are similar to many previous ESM studies (e.g. [9, 12, 37, 41, 51]), the generalizability of our findings to other groups like iOS users, the elderly, or individuals from other regions remains questionable due to possible differences in notification usage behaviors. In addition, our study was conducted during the school break. Therefore, for students the types and experiences of notifications might differ from regular days.

8 CONCLUSION

This paper, inspired by existing literature advocating for various notification management features, investigates how users utilize three such features—pinning, categorizing, and sorting—in their day-to-day lives. Our study provides insights into the actual usage, perceptions, experiences, strategies, and needs associated with these features. Results reveal that pinning was highly valued and utilized for its broad applicability in notification management. Participants also developed a new workflow—pin to temporarily preserve, bulk remove others, and then assess the pinned notifications—to efficiently manage their notifications. Regarding the other two features, results show how users utilized manual sorting to organize notifications and expressed a desire for automatic features to facilitate batch operations and processing. They also wanted these features to take their input to improve results and to adapt outcomes to activity contexts. Thus, this study contributes valuable empirical evidence to notification research, offering design implications for the development of more effective notification management systems.

ACKNOWLEDGMENTS

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 SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. Association for Computing Machinery, New York, NY, USA, 271–278. https://doi.org/10.1145/985692.985727

[2] Samaneh Aminikhanghahi, Ramin Fallahzadeh, Matthew Sawyer, Diane J Cook, and Lawrence B Holder. 2017. Thyme: Improving smartphone prompt timing through activity awareness. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 315–322.

[3] Christoph Anderson, Isabel Hübener, Ann-Kathrin Seipp, Sandra Ohly, Klaus David, and Veljko Pejovic. 2018. A Survey of Attention Management Systems in Ubiquitous Computing Environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2, Article 58 (Jul. 2018), 27 pages. https://doi.org/10.1145/3214261

[4] Julie Aranda, Noor Ali-Hasan, and Safia Baig. 2016. I'm just trying to survive: an ethnographic look at mobile notifications and attention management. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 564–574. https://doi.org/10.1145/2957265.2957274

[5] Nikola Banovic, Fanny Chevalier, Tovi Grossman, and George Fitzmaurice. 2012. Triggering triggers and burying barriers to customizing software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2717–2726. https://doi.org/10.1145/2207676.2208666

[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[7] Leslie H Brown, Timothy Strauman, Neus Barrantes-Vidal, Paul J Silvia, and Thomas R Kwapil. 2011. An experience-sampling study of depressive symptoms and their social context. *The Journal of nervous and mental disease* 199, 6 (2011), 403–409.

[8] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. arXiv:2005.14165 [cs.CL]

[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. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 1, Article 15 (Mar. 2020), 38 pages. https://doi.org/10.1145/3381000

[10] Chung Chiao Chang, Meng-Hsin Wu, Yu-Jen Lee, XiJing Chang, and Yung-Ju Chang. 2021. Opportune Moments for the Multi-Stage Notification Responding Process: A Preliminary Investigation. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers (UbiComp/ISWC '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 9–10. https://doi.org/10.1145/3460418.3479278

[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 (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 650, 17 pages. https://doi.org/10.1145/3544548.3581146

[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 (MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3338286.3340125

[13] 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 (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 6–15. https://doi.org/10.1145/2785830.2785852

[14] 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 (CHI '22)*. Association for Computing Machinery, New

York, NY, USA, Article 186, 18 pages. https://doi.org/10.1145/3491102.3517529

[15] Xinshi Chen, Shuang Li, Hui Li, Shaohua Jiang, Yuan Qi, and Le Song. 2019. Generative Adversarial User Model for Reinforcement Learning Based Recommendation System. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, 1052–1061. https://proceedings.mlr.press/v97/chen19f.html

[16] Yi-Shyuan Chiang, Ruei-Che Chang, Yi-Lin Chuang, Shih-Ya Chou, Hao-Ping Lee, I-Ju Lin, Jian-Hua Jiang Chen, and Yung-Ju Chang. 2020. Exploring the Design Space of User-System Communication for Smart-home Routine Assistants. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376501

[17] 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 *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 526, 15 pages. https://doi.org/10.1145/3491102.3517496

[18] Sunny Consolvo and Miriam Walker. 2003. Using the experience sampling method to evaluate ubicomp applications. *IEEE pervasive computing* 2, 2 (2003), 24–31.

[19] Fulvio Corno, Luigi De Russis, and Alberto Roffarello. 2018. AwareNotifications: Multi-Device Semantic Notification Handling with User-Defined Preferences. *Journal of Ambient Intelligence and Smart Environments* 10 (Aug. 2018). https://doi.org/10.3233/AIS-180492

[20] Mihaly Csikszentmihalyi and Reed Larson. 2014. Validity and reliability of the experience-sampling method. *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi* (2014), 35–54.

[21] 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

[22] Jon D. Elhai, Dmitri Rozgonjuk, Ahmad M. Alghraibeh, and Haibo Yang. 2021. Disrupted Daily Activities From Interruptive Smartphone Notifications: Relations With Depression and Anxiety Severity and the Mediating Role of Boredom Proneness. *Social Science Computer Review* 39, 1 (Feb. 2021), 20–37. https://doi.org/10.1177/0894439319858008 Publisher: SAGE Publications Inc.

[23] 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 (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering)*, Kazuya Murao, Ren Ohmura, Sozo Inoue, and Yusuke Gotoh (Eds.). Springer International Publishing, Cham, 74–89. https://doi.org/10.1007/978-3-319-90740-6_5

[24] 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)*. Association for Computing Machinery, New York, NY, USA, 181–190. https://doi.org/10.1145/2037373.2037402

[25] 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)*. Association for Computing Machinery, New York, NY, USA, 103–112. https://doi.org/10.1145/1851600.1851620

[26] Nicholas Fitz, Kostadin Kushlev, Ranjan Jagannathan, Terrel Lewis, Devang Paliwal, and Dan Ariely. 2019. Batching smartphone notifications can improve well-being. *Computers in Human Behavior* 101 (Dec. 2019), 84–94. https://doi.org/10.1016/j.chb.2019.07.016

[27] Sophia Hadash, Martijn C. Willemsen, Chris Snijders, and Wijnand A. IJsselsteijn. 2022. Improving understandability of feature contributions in model-agnostic explainable AI tools. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 487, 9 pages. https://doi.org/10.1145/3491102.3517650

[28] Guhyun Han, Jaehun Jung, Young-Ho Kim, and Jinwook Seo. 2023. DataHalo: A Customizable Notification Visualization System for Personalized and Longitudinal Interactions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 648, 21 pages. https://doi.org/10.1145/3544548.3580828

[29] Yuichi Inagaki, Ryoichi Shinkuma, Takehiro Sato, and Eiji Oki. 2019. Prioritization of Mobile IoT Data Transmission Based on Data Importance Extracted From Machine Learning Model. *IEEE Access* 7 (2019), 93611–93620. https://doi.org/10.1109/ACCESS.2019.2928216

[30] Shamsi T. Iqbal and Brian P. Bailey. 2008. Effects of intelligent notification management on users and their tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/1357054.1357070

[31] Shamsi T. Iqbal and Eric Horvitz. 2010. Notifications and awareness: a field study of alert usage and preferences. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 27–30. https://doi.org/10.1145/1718918.1718926

[32] Izabelle F Janzen and Joanna McGrenere. 2022. Reflective Spring Cleaning: Using Personal Informatics to Support Infrequent Notification Personalization. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing

Machinery, New York, NY, USA, Article 179, 16 pages. https://doi.org/10.1145/3491102.3517493
[33] Kiana Kheiri and Hamid Karimi. 2023. SentimentGPT: Exploiting GPT for Advanced Sentiment Analysis and its Departure from Current Machine Learning. arXiv:2307.10234 [cs.CL]
[34] Sunnie S. Y. Kim, Elizabeth Anne Watkins, Olga Russakovsky, Ruth Fong, and Andrés Monroy-Hernández. 2023. "Help Me Help the AI": Understanding How Explainability Can Support Human-AI Interaction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 250, 17 pages. https://doi.org/10.1145/3544548.3581001
[35] 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 (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1011–1020. https://doi.org/10.1145/2858036.2858359
[36] Alexandra Kuznetsova, Per B Brockhoff, and Rune HB Christensen. 2017. lmerTest package: tests in linear mixed effects models. *Journal of statistical software* 82 (2017), 1–26.
[37] 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
[38] 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 (MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article 40, 6 pages. https://doi.org/10.1145/3338286.3344387
[39] 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.
[40] Luis Leiva, Matthias Böhmer, Sven Gehring, and Antonio Krüger. 2012. Back to the app: the costs of mobile application interruptions. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services (MobileHCI '12)*. Association for Computing Machinery, New York, NY, USA, 291–294. https://doi.org/10.1145/2371574.2371617
[41] 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 Trans. Comput.-Hum. Interact.* 29, 5, Article 39 (Jan. 2023), 33 pages. https://doi.org/10.1145/3478868
[42] Elad Liebman, Maytal Saar-Tsechansky, and Peter Stone. 2015. DJ-MC: A Reinforcement-Learning Agent for Music Playlist Recommendation. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (Istanbul, Turkey) *(AAMAS '15)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 591–599.
[43] 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 (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 520, 13 pages. https://doi.org/10.1145/3411764.3445384
[44] Yong-Han Lin, Li-Ting Su, Uei-Dar Chen, Peng-Jui Wang, Yi-Chi Lee, and Yung-Ju Chang. 2023. Automatic, Manual, or Hybrid? A Preliminary Investigation of Users' Perception of Features for Supporting Notification Management *(UbiComp/ISWC '23 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 98–102. https://doi.org/10.1145/3594739.3610699
[45] Afra Mashhadi, Akhil Mathur, and Fahim Kawsar. 2014. The myth of subtle notifications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 111–114. https://doi.org/10.1145/2638728.2638759
[46] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (Nov. 2019), 23 pages. https://doi.org/10.1145/3359174
[47] 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
[48] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. 2019. NotifyMeHere: Intelligent Notification Delivery in Multi-Device Environments. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (CHIIR '19)*. Association for Computing Machinery, New York, NY, USA, 103–111. https://doi.org/10.1145/3295750.3298932
[49] 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)*. Association for Computing Machinery, New York, NY, USA, 813–824. https://doi.org/10.1145/2750858.2807544
[50] 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 (CHI '16)*.

Association for Computing Machinery, New York, NY, USA, 1021–1032. https://doi.org/10.1145/2858036.2858566

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

[52] Tadashi Okoshi, Julian Ramos, Hiroki Nozaki, Jin Nakazawa, Anind K. Dey, and Hideyuki Tokuda. 2015. Reducing users' perceived mental effort due to interruptive notifications in multi-device mobile environments. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 475–486. https://doi.org/10.1145/2750858.2807517

[53] OpenAI. 2024. GPT-4 Technical Report. arXiv:2303.08774 [cs.CL]

[54] 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

[55] 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

[56] Martin Pielot, Bruno Cardoso, Kleomenis Katevas, Joan Serrà, Aleksandar Matic, and Nuria Oliver. 2017. Beyond Interruptibility: Predicting Opportune Moments to Engage Mobile Phone Users. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 91 (Sep. 2017), 25 pages. https://doi.org/10.1145/3130956

[57] 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)*. Association for Computing Machinery, New York, NY, USA, 233–242. https://doi.org/10.1145/2628363.2628364

[58] 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 (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3319–3328. https://doi.org/10.1145/2556288.2556973

[59] 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 (MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3229434.3229445

[60] 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.

[61] 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 (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3055–3064. https://doi.org/10.1145/2556288.2557189

[62] Prasanta Saikia, Ming Cheung, James She, and Soochang Park. 2017. Effectiveness of mobile notification delivery. In *2017 18th IEEE international conference on mobile data management (MDM)*. IEEE, 21–29.

[63] Katarzyna Stawarz, Anna L. Cox, and Ann Blandford. 2014. Don't forget your pill! designing effective medication reminder apps that support users' daily routines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2269–2278. https://doi.org/10.1145/2556288.2557079

[64] Eran Toch, Hadas Chassidim, and Tali Hatuka. 2020. Can you Turn it Off? The Spatial and Social Context of Mobile Disturbance. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 91 (Oct. 2020), 18 pages. https://doi.org/10.1145/3415162

[65] Amirsina Torfi, Rouzbeh A. Shirvani, Yaser Keneshloo, Nader Tavaf, and Edward A. Fox. 2021. Natural Language Processing Advancements By Deep Learning: A Survey. arXiv:2003.01200 [cs.CL]

[66] Fang-Ching Tseng, Zih-Yun Chiou, Ho-Hsuan Chuang, Li-Ting Su, Yong-Han Lin, Yu-Rou Lin, Yi-Chi Lee, Peng-Jui Wang, Uei-Dar Chen, and Yung-Ju Chang. 2023. Multiple Device Users' Actual and Ideal Cross-Device Usage for Multi-Stage Notification-Interactions: An ESM Study Addressing the Usage Gap and Impacts of Device Context. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 649, 15 pages. https://doi.org/10.1145/3544548.3580731

[67] Liam D. Turner. 2017. *Decomposing responses to mobile notifications*. PhD. Cardiff University. https://orca.cardiff.ac.uk/id/eprint/102135/

[68] Liam D. Turner, Stuart M. Allen, and Roger M. Whitaker. 2015. Push or Delay? Decomposing Smartphone Notification Response Behaviour. In *Human Behavior Understanding (Lecture Notes in Computer Science)*, Albert Ali Salah, Ben J.A. Kröse, and Diane J. Cook (Eds.). Springer International Publishing, Cham, 69–83. https://doi.org/10.1007/978-3-319-24195-1_6

[69] 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.

[70] Aku Visuri, Niels van Berkel, Jorge Goncalves, Reza Rawassizadeh, Denzil Ferreira, and Vassilis Kostakos. 2021. Understanding usage style transformation during long-term smartwatch use. *Personal and Ubiquitous Computing* 25, 3 (June 2021), 535–549. https://doi.org/10.1007/s00779-020-01511-2

[71] Alexandra Voit, Dominik Weber, and Niels Henze. 2018. Qualitative Investigation of Multi-Device Notifications. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp '18)*. Association for Computing Machinery, New York, NY, USA, 1263–1270. https://doi.org/10.1145/3267305.3274117

[72] Zengzhi Wang, Qiming Xie, Yi Feng, Zixiang Ding, Zinong Yang, and Rui Xia. 2024. Is ChatGPT a Good Sentiment Analyzer? A Preliminary Study. arXiv:2304.04339 [cs.CL]

[73] Dominik Weber, Alexandra Voit, Jonas Auda, Stefan Schneegass, and Niels Henze. 2018. Snooze! investigating the user-defined deferral of mobile notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3229434.3229436

[74] Dominik Weber, Alexandra Voit, and Niels Henze. 2019. Clear All: A Large-Scale Observational Study on Mobile Notification Drawers. In *Proceedings of Mensch und Computer 2019 (MuC'19)*. Association for Computing Machinery, New York, NY, USA, 361–372. https://doi.org/10.1145/3340764.3340765

[75] 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

[76] Dominik Weber, Alexandra Voit, Philipp Kratzer, and Niels Henze. 2016. In-situ investigation of notifications in multi-device environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 1259–1264. https://doi.org/10.1145/2971648.2971732

[77] 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 (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3411764.3445209

[78] Dhanvanth Reddy Yerramreddy, Jayasurya Marasani, Ponnuru Sathwik Venkata Gowtham, S Abhishek, and Anjali. 2023. An Empirical Analysis of Topic Categorization Using PaLM, GPT and BERT Models. In *2023 Innovations in Power and Advanced Computing Technologies (i-PACT)*. 1–6. https://doi.org/10.1109/i-PACT58649.2023.10434768

[79] 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 (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5346–5360. https://doi.org/10.1145/3025453.3025946

[80] Lixin Zou, Long Xia, Zhuoye Ding, Jiaxing Song, Weidong Liu, and Dawei Yin. 2019. Reinforcement Learning to Optimize Long-term User Engagement in Recommender Systems. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA) *(KDD '19)*. Association for Computing Machinery, New York, NY, USA, 2810–2818. https://doi.org/10.1145/3292500.3330668

9 APPENDIX

A NOTIFICATION TYPE

Figure 7 depicts the types of notifications remaining after data cleaning, as described in Section 4.5. More than half (53%) of the notifications were from the Instant Messaging (IM) type, followed by Social with 12%, and then Tool with 9%.

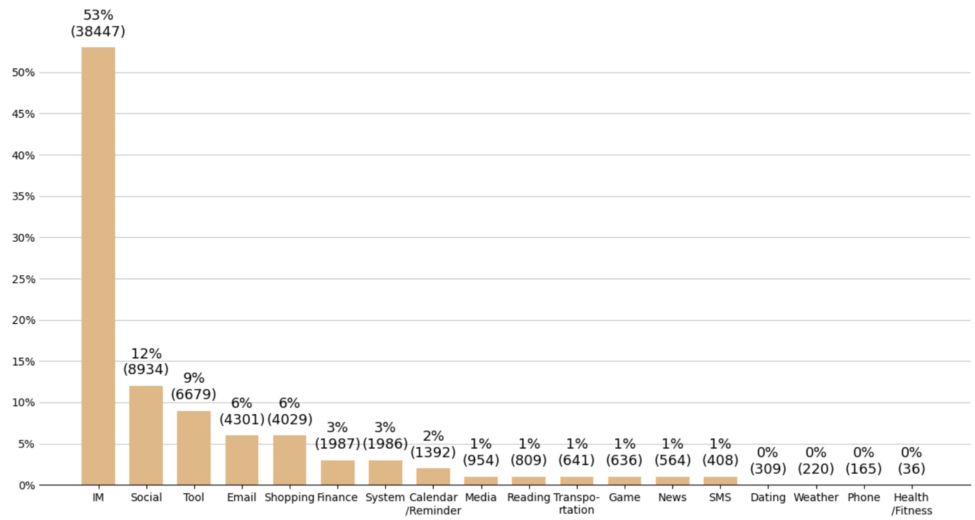


Figure 7. The percentage and number of notification types used in our study.