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

ANONYMOUS AUTHOR(S)

With the increasing number of smartphone notifications, users expend more effort in reading and acting on them. While prior research has proposed several notification management features, empirical evidence about these functionalities has been lacking. We developed a notification management app based on previous studies, encompassing sorting, and categorization, with both automatic and manual approaches, and manual pinning. Through a mixed-methods approach, this paper assesses the usage and perception of these features and compares their helpfulness across various approaches. Our results showed that among these features, pinning emerges as the most frequently used and the most helpful as it allows brief notification deferral, eases information access, and prevents accidental notification removal. We also found that user preference for features varies across different modes. Manual sorting aligns better with users' ideal notification order compared to automated sorting. In contrast, participants rarely created manual categories and used them less often than automatic ones for accessing notifications. Finally, we discussed the reasons and context for using and not using the features, as well as the factors influencing the preference for manual or automatic features.

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:
Anonymous Author(s). 2023. Pinning, Sorting, and Categorizing Notifications: A Mixed-methods Usage and Experience Study of Mobile Notification-management Features. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 0, 0, Article 0 (November 2023), 24 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Proliferating arrays of mobile apps are generating ever more diverse forms of notifications. This is challenging not only at the level of sheer quantity but also because existing smartphone interfaces and functionalities are inadequate to users' increasingly complex notification-management needs [24, 35, 44, 57]. The specific issue that has arguably attracted the most prior research attention is the disruptive/disturbing nature of notifications' sounds and vibrations [10, 26, 28], and recent proposed solutions have generally involved either automatic identification or user designation of suitable future times for notification delivery [1, 6, 20, 38, 40, 41, 43, 44, 56]. However, the problem of notification overload can only be partially addressed by these 'delaying tactics'.

Another common but less-discussed problem is that, amid the above-mentioned rapid growth in the volume of notifications mobile users receive, new notification-management needs and behaviors necessitate new smartphone features [24, 27, 35, 44, 57, 58]. For example, it was recently reported that the order in which notifications are presented on current smartphone systems does not align with users' preferred order [35]. Variation in users' definitions of notifications' roles, as well as how such variation leads to divergence in their notification-handling methods has also been highlighted [8, 34, 35, 46, 57, 58]. That is, a given notification may need to be saved,

2474-9567/2023/11-ART0 $15.00
https://doi.org/XXXXXXX.XXXXXXX

postponed, or used as a reminder; and indeed, some appear to require better methods of categorization than are currently available [24, 35, 56–58]. However, relatively few of the features that researchers have proposed to meet these divergent user needs have been empirically evaluated. Thus, it remains unclear if they can underpin major improvement in notification-management practices, or whether they are likely to be utilized often or only rarely. 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, categorizing, and pinning [35, 39, 56–58]. 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), 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 often, if at all, will users utilize sorting, categorization, and pinning in a notification interface when it is made available to them?
- RQ2: What are users' perceptions of the usefulness of these three features, separately and collectively?
- RQ3: Do users find manual or automatic sorting and categorization more helpful?

To answer them, we asked 30 participants to engage with all three features and three interaction modes in their day-to-day lives. More specifically, our Android research app had three distinct versions: the Manual, in which all features were controlled manually; the Auto, in which the features other than pinning were automated; and the Hybrid, which combined the other two modes. We then proceeded with a three-week within-subjects field evaluation, during which each participant interacted with each of these three app versions for a week apiece, and were asked to make comparisons among them.

This paper makes the following four main contributions:

- The study identifies pinning as the most commonly used feature among the three evaluated options. Moreover, it demonstrates that the pinning feature serves multiple purposes: not just deferring notifications but also preventing accidental deletion and facilitating future information retrieval.
- The study indicates a low frequency of manual categorization for organizing and accessing notifications. In contrast, automatically generated categories were more prevalently used by participants for notification navigation.
- The study reveals that automatically sorted notifications were not perceived to be more closely aligned with participants' desired notification order than the original, unsorted sequences.
- The research provides a nuanced understanding of why and how participants chose to use — or not use — the features of pinning, sorting, and categorization in their day-to-day notification management.

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 [11, 22, 41, 47] or enhance user engagement [2, 19, 42, 43], or controlling notification alerts to strike a balance between user awareness and potential disturbances [10, 36, 44]. 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 [36, 57]. 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 [10, 16, 30, 31, 36, 44, 45, 61]. Attending ranges from brief examinations of notifications to focused and committed engagement with them [16]. Quick interactions are typically measured by users unlocking their phones and viewing notifications [10, 36, 44], whereas longer engagement is measured by actions like tapping or dismissing them [10].

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" [9, 25, 49]. However, some scholars consider phone-unlocking and notification-tapping to attendance [10], 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" [7, 52, 54] for committed interaction, alongside other such terms including "seeing" [46], "focusing" [54], and "engaging" [7, 53]. 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 [7, 52]. 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, 12, 32]. 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 [44, 46, 48]. Yet, it's worth noting that previous studies have shown users don't always promptly respond to instant messages [12, 16, 32]. 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 [50, 56]. 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. [56]. Moving beyond just snoozing notifications, prior research has proposed an approach that allowed users to manually reposition notifications [35]. 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 [24, 35, 57]. 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, while previous research on notification management has offered a variety of feature recommendations aimed at assisting mobile users, these suggestions often lack empirical support to validate their real-world efficacy. Many of the discussed benefits remain theoretical, with no concrete evidence to confirm whether these features genuinely help users or how they are specifically utilized. Weber et al. [56] 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 aims to address this gap in the literature by providing empirical evidence on how smartphone users engage with these features in their daily lives. We delve into the specific reasons that drive both the use and non-use of these features, providing valuable insights for 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 behavior 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 offers functionalities including sorting, categorizing, and pinning notifications. NotiManager monitors and captures incoming notifications through the utilization of Android Notification Listener Service API [1]. All notifications, with the exception of ongoing ones (e.g. Google Navigation, media playback notifications), are 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 access and interact with them just as usual. 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:** We implemented the sorting feature in two modes, automatic and manual. In automatic sorting, notification order 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. [35], where participants in their study rearranged notifications arriving at their phone. 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 context factors like activities detected by the mobile sensors and the time when notifications were viewed. The sorted results are then sent back from the remote server and replaced with the original order, presenting a newly arranged display. 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 hybrid mode, a mode that combines all the features present in automatic and 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 for users to categorize notifications or process them within category, also available in both automatic and manual modes. Automatic categorization operates by classifying notifications based on app categories established by the Google Play Store and as suggested by prior research [48, 56, 57]. 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 delete-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,

[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

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

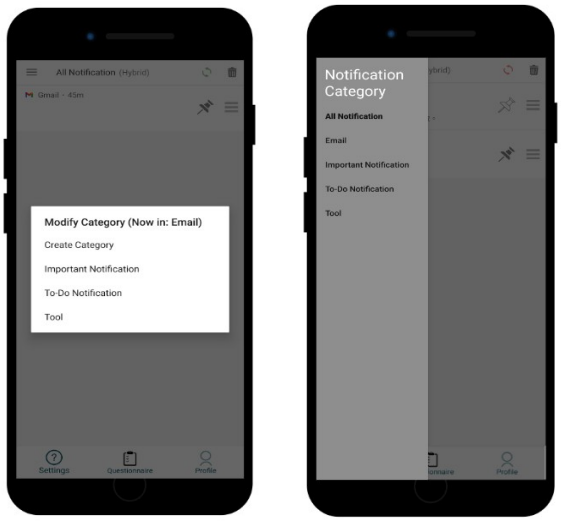
fulfilling a need identified in the literature [35] for a method to remind users to revisit notifications or to undertake specific tasks related to the notifications.

3.2 User Interface

The interface varies in different modes. The all-notification interface in 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 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 delete-all-notifications button (see Figure 1c). The 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 (see Figure 2a). 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 button (see Figure 1a) on the all-notification interface, the category menu (see Figure 2b) 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



(a) (b)

Figure 2. Modifying notification's category (a) and viewing category (b)

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 EVALUATION

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 Conditions

Participants were divided into either the manual group or the auto group, utilizing the default, manual, and hybrid mode, or the default, automatic, and hybrid mode during the three-week study period. 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.

To ensure participants were familiar with viewing and attending notifications through NotiManager during the study, we had each participant's initial mode set to the default mode for the first week. This mode replicated the original notification order, which placed IM and email notifications toward the top of the order. After the default mode, in the remaining two weeks of the study, each participant experienced two additional modes, one in each week: hybrid mode and either automatic or 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 automatic 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, or

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

manually adding new categories. In the hybrid mode, all functionalities from both the automatic and manual modes are integrated.

4.2 ESM and Diary Study

The aim of the ESM study [55] was to gain insights into participants' perceptions of the notification order under different modes and their motivations for managing notifications using NotiManager, considering various contexts. While the Diary study aimed to uncover participants' daily experiences and feelings while using these features over the course of a 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 in specific time, and logged phone-sensor data.

4.2.2 ESM Questionnaire. Each page of the ESM questionnaire was structured into several parts (see Figure 3a). Within different modes, the ESM questions vary. We administered questionnaires in the default mode to help participants become familiar with how to complete questionnaires by asking them about the extent to which the notification order at a specific time point matched their ideal sequence. To answer questions related to degree, our questionnaires all utilized a five-point Likert-scale. Respondents were also asked to rate the positioning of urgent and important notifications since prior work suggests that users prioritize notifications that are considered important or urgent [58]. We aimed to understand the context in which respondents encountered these sequences by asking their location [18, 59], ongoing activity [14, 17, 37, 39], their level of engagement in that activity [33], whether they were interacting with others at the time [6, 41, 51], and their degree of busyness during that moment [62].

In the automatic mode, respondents were additionally presented with comparative questions regarding two notification orders to evaluate our automatic sorting feature (see Figure 3b). Respondents were required to select the one closer to their ideal order. To address the potential difficulty in choosing, we introduced a third option: "Both orders are similar", which was made based on feedback from our pilot study. These two orders include the following distinctions: one represents the order in the current notification drawer, where IM and email notifications are positioned towards the top [57]. The second order is based on the predictions generated by our model. It is essential to highlight that the presentation sequence of the two notification display orders was randomized to minimize potential bias, as respondents were unaware of which order was the original and which one was the adjusted version.

In the manual mode questionnaire, when participants were not utilizing any management features, they were asked about the alignment of the current notification order with their ideal sequence. When management features were employed, respondents were questioned about their motivation, the perceived necessity of using them, how they ranked the importance and urgency of notifications addressed by the feature, and the context of feature use. Regarding the motivation question, we instructed participants to select only one option that best described.

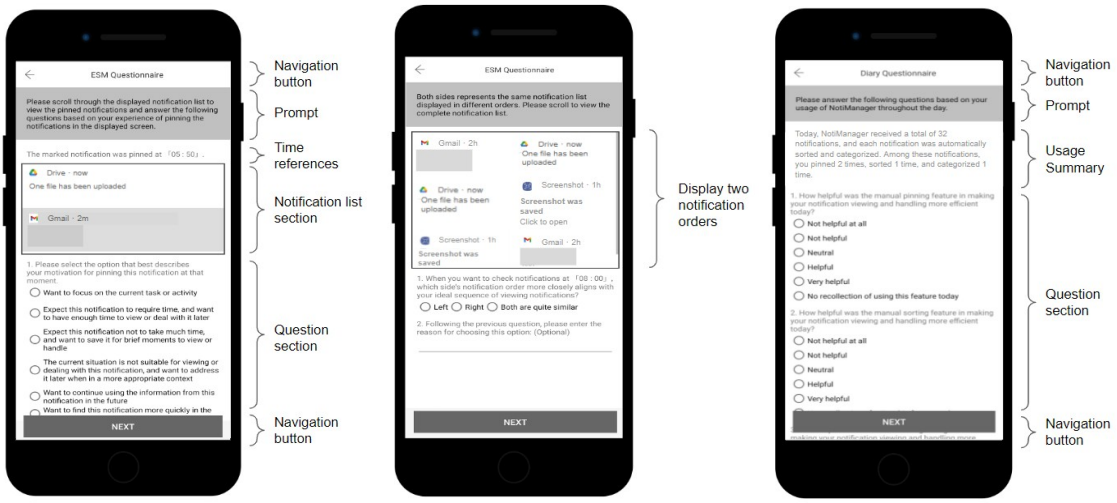


Figure 3. The ESM questionnaire interface (a), an example of ESM about comparative questions regarding two notification orders (b) and the Diary questionnaire interface (c)

The options included: 1) To focus on the current task or activity; 2) To defer notificaiton 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. These motivations were mostly adopted from Chang et al.'s study [8]. Additional motivations for categorizing function were: "To categorize notification 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 motivations were identified through our pilot study and personal use of NotiManager.

The hybrid mode questionnaire includes both automatic and manual mode question types. In each mode, we limit each questionnaire to address a maximum of three features. When selecting which three features to include, we prioritized diversity, ensuring that each utilized feature type was represented at least once.

4.2.3 Diary Questionnaire. Each page of the Diary questionnaire was organized with serveral components (see Figure 3c). In the Diary questionnaire, we employed a five-point scale to inquire about the perceived helpfulness of each feature. Following the questions on the helpfulness of these features, a text-box question was also provided for respondents to convey their overall feelings and insights regarding their experiences throughout the day of using NotiManager.

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 changes in the displayed notification order occurred within the last 30 minutes, a time threshold derived from prior research (e.g., [9]). 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, on the other hand, 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. All three pilot participants could complete the ESM and Diary questionnaire within an average of three minutes, an acceptable duration for an ESM study [5, 13, 15], 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. Over the three-week study, participants encountered a distinct version of NotiManager every week. The version is automatically switched after a week and we sent an email to notify participants of the upcoming automatic switch to the new version. Within NotiManager, a pop-up window was also shown to indicate that the version had been updated. Toward the study's end, we invited participants to semi-structured interviews by email, and 28 individuals partook. These interviews aimed to delve into various aspects, including participants' feature preferences and perspectives, their notification management approaches in the original notification drawer and NotiMangaer, and details about their ESM and Diary responses. Compensation was based on the number of completed ESM and Diary questionnaires, at US$0.3 each. 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. These criteria 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. All participants reported receiving an average of more than 10 notifications per day and using their phones for over 2 hours daily.

4.5 Quantitative Data Cleaning and Analysis

We recorded 109253 notifications in total and received 2090 ESM responses along with 506 Diary responses, with a response rate of 71% and 80%, respectively. We excluded 36756 notifications that only appeared in the default mode, which is a mode primarily intended for participants to become accustomed to using NotiManager. This left us with a dataset comprising 72,497 notifications from 424 unique apps. The majority of these notifications were dispatched by three leading messaging apps: *LINE* (21,336 notifications; 27 participants), *Telegram* (8,566 notifications; 6 participants), and *FB Messenger* (6,823 notifications; 26 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 manual and automatic modes and to

gauge the perceived helpfulness of individual features, we conducted a logistic regression analysis with the "lmerTest" [29] package in R software[3], 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 [4], an online collaborative qualitative-analysis software. We employed an inductive thematic analysis approach [4], wherein the development of the codebook was an iterative process. Initial codes were generated based on the researchers' notes from three interviews, along with memos that explained the meanings of the codes. 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 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 in the emergence of numerous themes that pertain to how different management features influenced participants' behaviors and perceptions, including in manual and automatic mode.

5 QUANTITATIVE RESULTS

We begin by providing a summary of using NotiManager and the three features. Subsequently, we delve into how each specific feature was utilized and participants' sentiments regarding the effects brought about by each function, such as their perceived level of helpfulness.

5.1 Usage Overview

Each participant entered NotiManager an average of 27 times per day (Md = 21; SD = 20.91). On average, each of them spent 8 seconds from entering to leaving NotiManager, with the shortest duration being less than 1 second and the longest lasting 1949 seconds. Within NotiManager, pinning was utilized the most (230) among the manual features, followed by manual sorting (162), and manual categorization (34). Regarding perceived helpfulness, which is rated on a five-point scale, we observed a similar order: pinning was also rated as the most

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

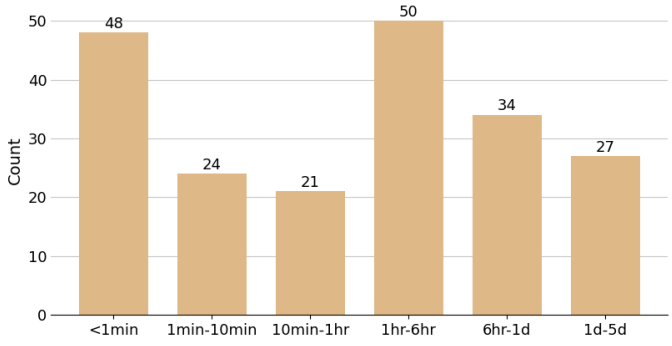


Figure 4. Distribution of Participants' Pinning Durations

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).

5.2 Pinning Notifications

Out of the 30 participants, 22 engaged in the act of pinning notifications. We logged a total of 230 pinning events, involving 219 unique notifications that spanned across 27 different apps. Notifications related to communication dominated the pinning activity, with 57% originating from IM applications. The primary contributors were *LINE* (97 notifications), *Gmail* (25 notifications), and *FB Messenger* (23 notifications). When examining instances in which participants accessed NotiManager and performed at least one pinning action, the maximum number of pinning events in a single session was 6, with a mean of 1.44 (SD = 0.87). The duration for which notifications were pinned is presented in Figure 4. The average duration of a pinned notification was 564.6 minutes (SD = 1103.24). Interestingly, a quarter of all pinned notifications (24%) were unpinned within just one minute, while about one-third (35%) had a pinning duration of less than ten minutes. About half of all pinning durations (46%) were under an hour, with a median duration of 80.45 minutes. A notable portion, 13%, remained pinned for more than one day, with the longest observed duration being nearly five days (7,105 minutes).

As depicted in Figure 5, the primary reason, accounting for 31%, was to defer addressing the notification. Participants often felt that the content of the notification would require a more substantial time commitment, and they chose to attend to it when their schedule permitted. Intriguingly, the second most cited reason was the desire to prevent accidental deletion of the notification, representing 27% of the responses—this percentage closely rivals the primary reason for deferral. The third most cited reason was to facilitate easier access to the information in the notification later, garnering 15% of the responses.

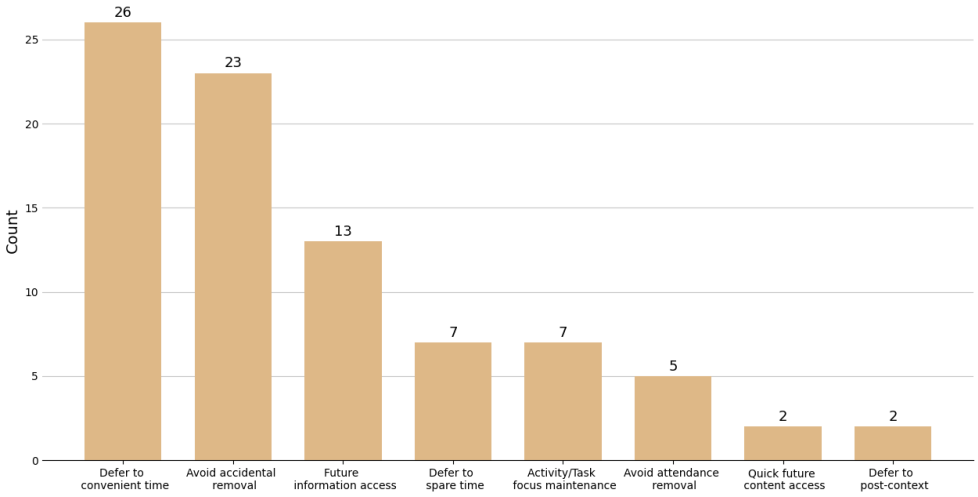


Figure 5. Participants' self-reported motivations for pinning notifications. (Only one motivation was provided for each pinning instance.)

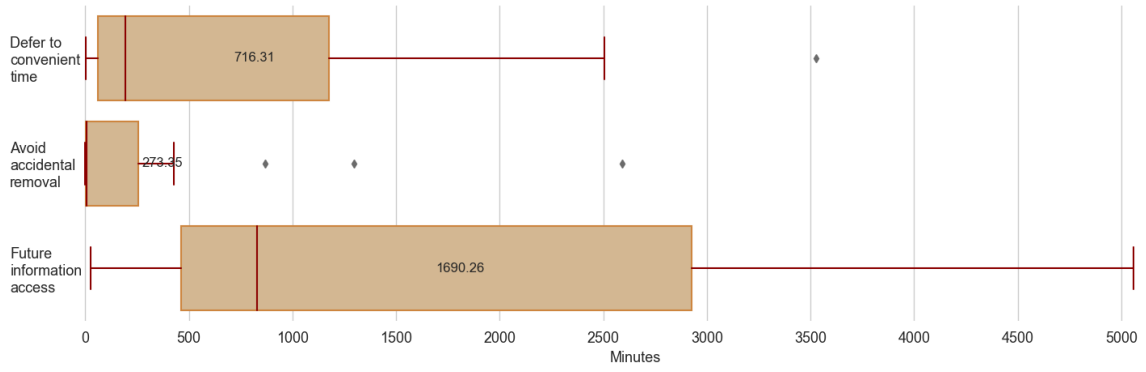


Figure 6. The duration of pinning corresponding to the top three selected pinning reasons.

The diversity in motivations for pinning notifications was further reflected in the variable durations for which notifications were pinned, as shown in Figure 6. When the primary reason for pinning was to facilitate future access to the information contained within the notification, the average duration of pinning was notably long (M = 1690.26, SD = 1634.51). The logistic regression analysis showed this was statistically longer than the durations observed when the primary reason for pinning was to protect the notification from inadvertent deletion ($t(65.96) = 2.188$, $p = 0.0322$, SEM = 367.4). The pinning duration of preventing accidental deletion was relatively brief, averaging just 273.35 minutes (SD = 586.52). However, when the primary reason for pinning was deferral, there is no statistically significant difference in durations compared to other reasons (vs. information access: $t(65.74) = 1.352$, $p = 0.18094$, SEM = 353.09; vs. avoid accidental deletion: $t(57.77) = 1.365$, $p = 0.1777$, SEM = 239.3). In addition, we observed intriguing behavior patterns associated with the reasons for pinning. Specifically, when participants pinned notifications to prevent accidental deletion, 83% of these events were subsequently followed by an action to delete all notifications. This was in stark contrast to when the reason was to ease future access to information, which has only 7.7% ($Z = -2.337$, $p = 0.0194$, SEM = 1.6301). When the reason was to defer the notification, 64% of these events were subsequently followed by an action to delete all notifications. The value was not significantly lower than when the reason was to prevent accidental deletion ($Z = -1.49$, $p = 0.1362$, SEM = 0.8396). In summary, our participants pinned notifications for various reasons, and the differences between these reasons were partly evident in both the pinning duration and subsequent actions taken with those notifications.

5.3 Sorting Notifications

Among the 30 participants, 22 participants made use of the manual sorting feature. A total of 162 manual sorting events were recorded from 140 distinct notifications spanning 40 different apps. Again, participants the most often adjusted the position of communication-related notifications, with the top three being *LINE* (39 notifications), followed by *FB Messenger* (19 notifications), and *Gmail* (16 notifications). The IM category compose 46% (64) of the all manual sorting notifications. Notably, nearly one-fifth (19%) of the manual sorting events were on pinned notifications. When examining instances in which participants accessed NotiManager and performed at least one manual sorting action, the maximum count of manual sorting events was 3 (M = 1.4, SD = 0.6).

Figure 7 elucidates the primary reasons for manually adjusting the position of specific notifications among participants. The most prevalent reason, constituting nearly two-fifths (38%) of the instances, was to facilitate quick location of the notification at a later time. The second most common reason, comprising 19% of cases, was an expectation that the notification would not necessitate a significant time commitment, leading participants to

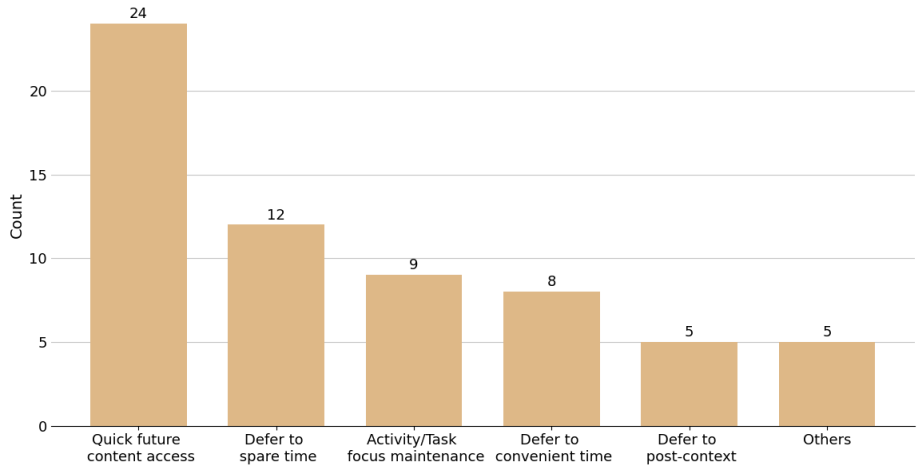


Figure 7. Participant's self-reported reasons for sorting notifications

prefer addressing it during subsequent brief moments of free time. This was followed by the desire to maintain focus on their current task or activity, a reason cited in 14% of cases. The latter two reasons are both related to the concept of deferral but differ subtly in the underlying motivations for choosing to defer the notification.

In our evaluation of the automatic sorting feature, participants were asked to compare the sequence generated by our model against the original notification sequence via the ESM questionnaire. Out of a total of 864 comparison instances, the original notification sequence was selected 62.5% of the time (or 540 times). On the other hand, the automatically sorted notification list was chosen 119 times, and in 205 instances, participants considered both orders to be similar (see Figure 9). Regression analysis confirmed that the likelihood of choosing the original notification sequence over the automatically sorted one was statistically significant (Z = 2.162, p = 0.0306, SEM = 0.2225). Participants also assess whether a presented presentation order aligns with their ideal order. As shown in Figure 10, participants felt that notification orders after manual sorting were more in line with their ideal order (M = 3.26; SD = 0.82) than those in automatic (M = 3.01; SD = 0.87, t(493.6832) = -2.337, p = 0.0198, SEM = 0.1689). However, as Figure 8 illustrates, the diary data revealed that participants did not perceive a statistically significant difference in the level of helpfulness between manual and automatic sorting methods (Manual: M = 3.54, SD = 0.82, Automatic: M = 3.01, SD = 1.37, t(334.24776) = 0.633, p = 0.527, SEM = 0.07347).

5.4 Categorizing Notifications

In terms of category creation and usage in the hybrid and manual modes, our data showed minimal engagement with the feature of manual category creation. Specifically, only one participant (P29) engaged in creating a new category while operating in hybrid mode; none did so in manual mode. This participant initiated a total of two 'add-new-category' events and used the created category for notification viewing 3 times. However, although the act of manually creating new categories was rare among participants, the utilization of existing categories for classifying notifications and accessing notifications later was considerably more frequent. Specifically, a total of 31 instances were observed in which participants classified notifications using the pre-defined categories available in the system. Additionally, 277 instances were recorded where participants accessed notifications via

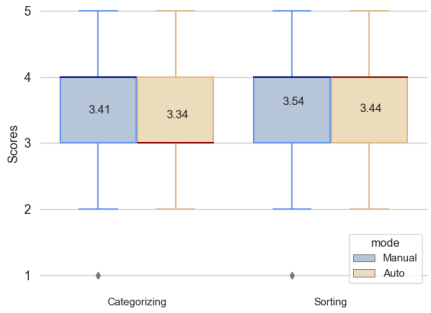


Figure 8. Participants's Perceived Helpfulness for Manual and Automatic Features (Measured on a 5-Point Scale)

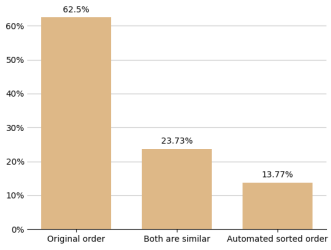


Figure 9. Participants' preference between the original notification order and the order after automatic sorting.

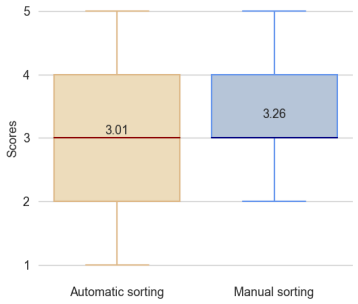


Figure 10. Participants' subjective ratings on how closely the notification order aligned with their ideal preference. (1=Not aligned at all; 5=Strongly aligned)

these existing categories, whether those categories were generated manually or automatically. The distribution of these 277 instances is clearly illustrated in Figure 11. A stark contrast is evident between the frequency of usage for automatically generated categories (274 instances) and manually created ones (3 instances). This disparity underscores the prevalent reliance on system-generated categories over manually created ones for managing notifications. In the instances where manual categorization was employed, *LINE* stood out with 6 notifications being manually categorized. Among the categories automatically generated, the most clicked category was *IM* (n=60), followed by *Social* (n=34) and *Transportation* (n=31).

In instances where participants accessed NotiManager and took at least one action to access notifications via a category, the frequency of such events varied, with the maximum count reaching up to 11 events (M = 2.69, SD = 2.29). Interestingly, although categorizing notifications was less frequent than pinning, participants occasionally used the clear-all feature within category page. Specifically, 6% of the clear-all events were executed within specific categories, as opposed to the main page where all notifications are displayed. Additionally, 25% of clicks to access a category led to the usage of the clear-all feature.

As for the motivations behind categorizing notifications, only a limited number of self-reported responses were received (n=8). As depicted in Figure 12, exactly half of these responses (n=4) indicated the intent to collectively process notifications that belong to the same category, aligning with our expectations. However, there were other reasons cited for categorizing notifications; due to the limited number of instances, these additional motives are not elaborated upon in this analysis.

Intriguingly, even though interactions with automatically generated categories were more prevalent among our participants, there was no significant difference in their perception of the helpfulness of manual versus automatic categorization. As Figure 8 illustrates, the mean scores for perceived helpfulness did not significantly differ between manual (M = 3.41, SD = 0.88) and automatic categorization (M = 3.34, SD = 0.82). Further regression analysis confirmed this lack of statistical significance ($t(297.3448) = 0.235$, $p = 0.815$, SEM = 0.0817).

6 QUALITATIVE FINDINGS

To address RQ2, we queried participants about when and how they found utility in employing these features. Echoing the quantitative findings discussed earlier, two significant takeaways emerged: firstly, the contextual

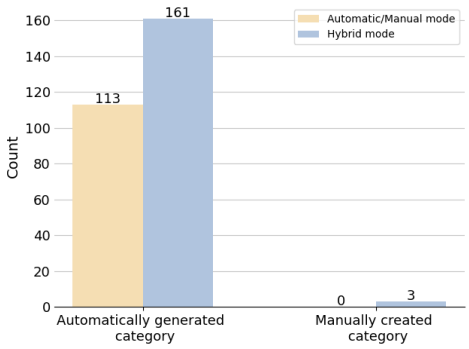


Figure 11. Participants' clicks on automatically generated vs. manually created categories across different modes.

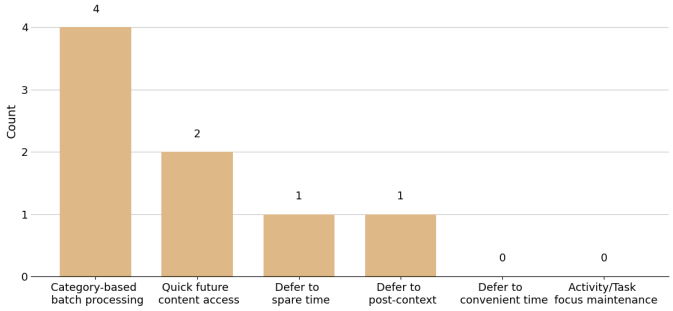


Figure 12. Participant's motivation for categorizing from the responses to the ESM questionnaires.

factors, rationales, and the ways in which users employ pinning; secondly, the reasons behind the underutilization of manual sorting and categorization.

6.1 Understanding the Context, Reasons, and Usage of Pinning

Most commonly, participants mentioned pinning notifications for those they needed to handle but could not at the moment due to their current context or because the notification required a substantial amount of time to process. They found it essential to keep these notifications undeleted to stay focused on their ongoing activities and to use them as reminders to ensure they do not forget to take action when they revisit the notifications. The types of notifications mentioned most frequently for this situation were IM, emails, online order pickup notifications, and longer content notifications. As P18 noted, he pinned notifications because he wanted to defer the action in the notification on another device afterward: *"It was a music company sending me an email asking me to download music files, and I pinned it. I wouldn't do it on my phone because I have a computer, speakers, and more professional equipment at home. So, I pinned it, and when I was sitting in front of my computer, I saw the pinned notification. I then opened the Gmail on my computer to download these things."* Furthermore, some participants mentioned that when they are unable to handle notifications immediately, they prefer to pin them to avoid accidental deletion. As noted by P14: *"If I pin a notification, it's probably because I have a notification but do not have the time to check it right now, and I intend to look at it later. Pinning it ensures that I do not accidentally dismiss it."* Similarly, in situations where they could not immediately act on a notification, P21 discussed pinning notifications that required more time to handle when she was busy or not in the mood to act upon them, but still wanted to check the content first: *"For example, in group chats for clubs, sometimes there were messages that seemed like they'll take a while to handle, but I wanted to see what the content is first. So, I pinned it and then clicked to see."*

In the realm of deferring notifications for later processing, the existing literature, as well as insights from some participants, have illuminated the concept of snoozing notifications. This feature allows notifications to pop out at a more opportune moment. To gain a deeper understanding, we sought the perspectives of participants on the difference between snoozing and pinning. Those inclined toward snooze functionality have underscored its capacity to reintroduce notifications without requiring direct engagement with their mobile devices. As explained by P20, *"I initially pinned the notification with the intention of postponing it. However, despite pinning it, [...] I completely forgot about the pinned notification. It was only when I revisited my notification panel that I rediscovered the task. In contrast, if I set a specific snooze period, [...] the notification resurfaces as I reach my destination, it*

acts as a helpful reminder." Conversely, individuals who preferred the pinning feature emphasize its simplicity of operation—a mere click preserves the notification. Furthermore, pinning offers greater flexibility in time management, particularly in cases where the exact timeframe for addressing a notification remains uncertain: *"If I use pinning, assuming I initially planned to have free time in two hours, but I become available after just one hour, I can go ahead and handle it directly. However, if I use snooze, I might forget in the meantime, and I may have to wait until the full two hours to address it."*(P10) Furthermore, several participants expressed concerns that snoozing notifications could create a sense of urgency or pressure, similar to alarms, which might be a bit stressful for them. As mentioned by P29, *"If it suddenly pops up after just 30 minutes, it feels like I'm being chased by someone."* Additionally, some participants have expressed concerns about snoozed notifications with alerts, potentially causing confusion when an alert sounds, leaving users uncertain whether it signals a new notification or a snoozed one. As P4 commented, *"I'm concerned that using snooze will make my messages very disorganized. I won't be sure when each notification was sent to me, and I might get confused or annoyed."* Hence, despite that several participants considered using the delay feature, which involves postponing notifications until an appropriate time, many more participants preferred to receive notifications instantly and pin them if found necessary.

The usage of pinning was combined with other features. As pinned notifications resist deletion, a subset of participants opted to combine this feature with one-click delete all. As explained by P2: *"Because deleting messages one by one is too slow, sometimes I just used the one-click delete all, which may accidentally delete important messages."* P15 echoed the situation that due to the high volume of daily notifications, he preferred to clear unnecessary ones first and then address the pinned notifications when he has more available time, *"I kept notifications [that required some time to address] pinned, allowing me to clear other notifications because I received numerous notifications throughout the day, and the pinned one won't be dismissed."* In P22's scenario, she used pinning and one-click delete all as initial steps in clearing notifications, *"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."* In addition to using one-click delete all in conjunction with pinning, participants sometimes employed manual sorting after pinning to facilitate the reordering of notifications. However, participants' preferences for the placement of pinned notifications varied considerably. For some, they usually manually moved pinned notifications to the top of the notification interface which is easier to notice, considering them as valuable visual reminders, as P9 stated, *"I prefer immediately seeing manually sorted pinned notifications at the top every time. I want each glance to serve as a reminder of their significance."*. In contrast, others had a more flexible approach, suggesting that pinned notifications could be interspersed with automatically sorted ones. For the participants, the act of safeguarding notifications through pinning held greater importance than their specific placement within the notification drawer, as stated by P12, *"My primary concern with pinning is to ensure the notification remains undeleted, irrespective of its position in the drawer."*. Intriguingly, a specific subset of participants favored situating pinned notifications at the bottom of the list. The underlying logic for this preference emanated from their belief that these notifications had already been examined, and therefore should not interfere with the visibility of unread messages. For example, Participant 23 expressed a desire to position pinned notifications at the lower end of her list. This organization facilitated her ability to swiftly discern that, upon encountering these pinned items, she had successfully reviewed all other unread notifications.

Nevertheless, there are still participants who do not utilize the pinning feature. Their reasons for not using pinning are typically that they don't see the need for preserving notifications that shouldn't be deleted. If there are notifications they want to keep, they simply leave them as they are. Alternatively, they do not engage in using one-click delete all, as P3 noted, *"Since I glanced at every notification and swiped away the ones I do not need, and I do not use the clear all option, pinning might not be as important for me."* Additionally, some individuals

mentioned that they prefer to rely on their own methods to remember to handle notifications. For instance, they may use their calendar (P1, P18) or the built-in features of the app (P13, P20) for this purpose.

6.2 Manual vs. Automatic: Comparing Sorting and Categorization Utilization

In line with the quantitative data, participants mentioned that they engage less in manual classification and sorting of notifications. They cited reasons including not wanting to invest too much time in managing notifications, such as the need to drag notifications for manual sorting or the additional steps required for categorization through extra clicks or manually adding notifications to categories. They feel that simply viewing all notifications without using these features is more efficient. As stated by P27, *"I believe notifications should be visible at first glance, and that's the most convenient approach for me. I do not want to expend extra effort to click into various categories."* Similarly, P28 said, *"I don't see a strong need for manual sorting, as I typically review each notification individually."* Additionally, several participants utilizing the Hybrid mode felt alternative features already provided comparable outcomes, diminishing the appeal of manual functions. Specifically, P14 and P20 believed that automatic categorization and sorting met their needs, deeming manual functions unnecessary, as P20 stated, *"I feel that the default categorization is already doing well, especially with categories like social notifications. I do not find a need to introduce more categories."* P14 shared a similar perspective, saying, *"Automatic sorting had put the notifications I want to see right up there at the top. It's the same as I do it myself."* Moreover, participants cited various reasons for not using manual sorting. Some were indifferent to the notification order, while others found the task of deciding how to sort notifications to be cumbersome.

For those participants who did engage with the manual features, they occasionally combined them with the pinning function we discussed earlier. Moreover, some mentioned that if a notification didn't appear where they initially expected it to be, but they wanted quicker access to it later, they would manually move the notification to their desired location. For instance, whereas P6 placed her boyfriend's chat messages at the top, P23 moved lab-related notifications to the bottom on a day when she did not have the mood to focus on work. In the case of P26, she arranged notifications she planned to view later towards the top for quick retrieval when needed, *"If I consider something important, I move it to the top, [...] I could access them right away when I open the app without the need for additional scrolling."*

The primary motivation for participants to engage in manual categorization was to correct inaccuracies in the automated categorization, as P14 explained: *"The automatic categories did not always match what I have had in mind. For instance, it had wrongly grouped a messaging app under 'Finance,' which was not where I would expect it to be."* When it comes to navigating through categories to find specific notifications, most participants used this feature when they had an abundance of notifications, and they wanted to access specific ones that might be buried deeper in the list. As noted by P13: *"There were times when I was so busy that checking every single notification was not an option. In those cases, I usually started by focusing on important communications, like messages from apps like LINE."*

7 DISCUSSION

7.1 How and Why did Participants Frequently Pinned Notifications

Our quantitative and qualitative results reveal that pinning was the feature most commonly used and preferred by the participants, and deferring responses until a more convenient time was their main reason for using it. In the interviews, we asked participants to evaluate the pros and cons of the snoozing and pinning features, particularly in terms of how they helped with deferring notifications. One of our key findings aligned with Weber et al.'s [56]: namely, that snoozing can sometimes be inconvenient due to the cognitive load of deciding upon a deferral duration and the effort required to input it. Our data also highlighted that pinning had comparative

advantages, notably 1) enabling the taking of quick, simple actions without such deliberations and 2) obviating the need for pre-scheduled re-engagement.

It was widely perceived that the snoozing feature's notification alerts, once they appeared, served as reminders. Interestingly, while some participants found this useful, others felt it created pressure, as though they were being "chased" by their tasks. Pinned notifications, in contrast, were seen as less intrusive because of their less active character; but among those participants who checked their phones frequently, the pinned items' constant presence in the notification interface nevertheless acted as subtle reminders, especially if positioned prominently within it.

Crucially, however, deferral was just one of several key reasons participants cited for pinning notifications. Others included enabling easier future access to the information contained within them and preventing their accidental removal. This multifaceted behavior pattern suggests that our participants often anticipated future scenarios in which notifications' information might be needed, e.g., for task completion. By keeping these notifications pinned in a readily visible and accessible location, they were trying to ensure a seamless pathway to retrieving that information when they needed it later.

Concern about unintentionally deleting notifications drove the major rationale for pinning. Pinning duration for this purpose was found to be significantly brief, but they nevertheless gave participants enough time to make preliminary assessments of whether a notification deserved further attention or action. Under existing notification-dismissal mechanisms, pinning after tapping notification is not possible, as it has already disappeared. Often, our participants' goal was merely to declutter their notification centers, but the sheer volume of its contents made it difficult for them to distinguish which notifications were important. In response to this challenge, some adopted a proactive approach: pinning notifications that seemed like they might be vital before viewing their content. This preemptive step ensured that such notifications would remain visible even after a bulk-removal feature was employed to clear the other, i.e., un-pinned notifications. Then, once nothing remained in their notification centers except pinned items, the participants felt free to carefully review their content at their leisure. As such, this practice served a dual purpose: i.e., it not only safeguarded important notifications but also provided a systematic approach to managing the influx of daily notifications. Notably, this "pinning before assessing" behavior highlights a problematic yet challenging-to-address feature of the current notification system: notifications are dismissed upon interaction. Interestingly, pinning served as an intuitive workaround, resolving concerns about accidental deletions and allowing for worry-free initial assessments of whether further action was needed. However, while the pinning feature seems to currently addresses this challenge, there is room for future systems to address it more seamlessly and intuitively. For example, future systems may incorporate features that allow for quick, risk-free assessment of notifications, and/or integrate capabilities that ensure that important information in notifications can be easily retrieved when needed.

Taken together, our data reveal that the act of pinning notifications is far from a one-dimensional behavior aimed merely at deferring attention. Instead, it serves a variety of user needs. Two main reasons for using the pinning feature – preventing accidental notification-deletion and enabling easier future access to notification content – distinguish it clearly from the simpler act of delaying or snoozing notifications. This rich diversity in pinning's usage suggests that, while there may be conceptual overlap between pinning and snoozing, they are not merely redundant features; the former addresses distinct user needs that the latter does not. Hence, we recommend that future systems incorporate both snoozing and pinning. Moreover, our findings show that the different reasons for pinning are associated with distinct durations for which the notifications remain pinned. This suggests the potential for real-world systems to distinguish between users' reasons for pinning based solely on pinning duration, eliminating the need for ESM to determine user intent.

7.2 Notification Management Features: Automatic vs. Manual Functions in Future Notification Systems

Although our study has not provided conclusive quantitative evidence in favor of either automatic or manual sorting/categorization of notifications, its findings do shed light on why the participants only used these functionalities infrequently. Below, we discuss these reasons and provide recommendations for future development of automatic vs. manual features.

7.2.1 Automatic vs. Manual Sorting. In some specific cases, our participants felt that the automated system accurately positioned the notifications they deemed important. However, their overall sentiment was that the automatically sorted sequences did not closely align with their preferred arrangements. This broadly unsatisfactory performance could have been linked to the limited features of our training model, which focused solely on the content and time stamp of notifications and neglected broader contextual information.

The training data themselves were sourced from a relatively small and distinct set of participants, which would have affected both the model's generalizability and its predictive ability. Moreover, it did not incorporate user behaviors like notification engagement or the outcomes of manual-sorting efforts. Therefore, its less-than-optimal performance was not entirely unexpected. Nevertheless, our results raise questions about the advisability of investing in automated-sorting features. Additionally, some participants who generally skimmed through their notifications expressed a lack of concern about their order, implying that they would process them swiftly irrespective of their arrangement. This, in turn, suggests their lack of motivation to evaluate notification sequences, which could also have contributed to the smallness of the observed difference in user preference between unsorted and automatically sorted sequences.

On the other hand, manual sorting was nearly one-fifth of the time used in combination with the pinning function. In the other cases, the participants chose to manually sort notifications also when they intended to preserve specific ones for future attention instead of immediately addressing them. The positions to which they moved these notifications varied depending on their intent. Specifically, moving notifications to the top of the stack was generally to enhance their visibility and thus their utility as reminders. On the other hand, relegating notifications to lower positions often indicated that the user's review of notifications was complete. Beyond this reason, which aligns with previous research findings [35], our study also uncovered an additional motivation for moving notifications to lower positions: to make more 'screen real estate' available for unread notifications.

In light of these observations, we recommend that the designers of future notification systems focus on manual-sorting capabilities, particularly to augment the pinning feature. Where automatic sorting is also considered for inclusion, perhaps on the pattern of our hybrid approach, we recommend that its models take account of a broader range of factors, including but not limited to user behaviors like clicking, pinning, and manual sorting.

7.2.2 Automatic vs. Manual Categorization. Among the 30 participants, only one had used the manual categorization function to add new categories, categorize notifications, and then navigate through their notifications using these new categories. Mostly, the participants felt that the time they would spend creating and sorting categories would equal or even surpass the time they would spend merely navigating through notifications if such categories did not exist. Moreover, most of our participants did not perceive a need for additional categories to manage their notifications, especially not when their systems already had an automatic categorization function; i.e., many believed that auto-generated categories were adequate, and this discouraged them from adding personalized ones. Therefore, although past researchers have recommended letting users customize notification categories in the same way that email software allows in the case of emails [57], our results imply that smartphone users feel relatively little incentive to do this. We speculate that this low motivation could stem from mobile users' general inclination to dismiss notifications as a (not always accurate) means of helping themselves feel that nothing important has been left undone [8, 57]. When they opt to retain notifications, it is primarily for the

reasons previously mentioned in the context of pinning: e.g., deferring engagement, ensuring swift access to information at a later time, and preventing unintentional deletion. Nonetheless, their end goal remains to remove these notifications sooner rather than later. This contrasts with emails, which can accumulate over months or years, necessitating more organizational features to streamline and categorize their content for efficient future retrieval [21, 23, 60].

Also, on the whole, our participants seldom used categorization to navigate through notifications. As many of them noted, most notifications only required a quick scan, taking minimal time, and therefore did not warrant extra viewing steps. The main occasions on which they found categorization useful involved having many notifications in the drawer and needing to find a specific type of them. In such cases, they found that auto-generated categories did help them locate specific types of notifications more efficiently, a result resonating with the suggestion for including such a feature from a recent study by Chang et al. [8]. Thus, automatic categorization can be valuable for users who frequently navigate though a high volume of notifications to find a specific type of notifications.

Notably, we did not use machine learning, but rather straightforward rules based on app-store categories, for which the categorization outcomes were still deemed sufficient for many participants' needs. Nevertheless, we would encourage future researchers and practitioners to explore more advanced notification classifiers that take account of factors like urgency, actionability, and specific content. Likewise based on the evidence from the current study, we do not recommend adding manual categorization features; however, future research should continue to explore the circumstances under which smartphone users might want to use it.

7.3 Study Limitations

The current study is subject to several limitations. First, our notification management features were unable to be implemented directly on the Android notification drawer. Thus, participants had to enter NotiManager to view and act on their notifications, which could affect their usage of the management features. Second, despite having a week-long default mode to familiarize with NotiManager, it might still not be sufficient for the participants to truly get accustomed to the functions, especially regarding the pinning and categorizing features. Not clicking precisely on the pin icon might lead to accidentally attending notification. Once a notification is attended to, it can no longer be pinned as it disappears from NotiManager. Moreover, categorizing is comparatively more complex to operate, potentially requiring more time for users to get used to it. Third, the sample size of ESM questionnaires related to each function was small although efforts were made to adjust the ESM sending mechanism through a pilot study. Fourth, during our interviews, we asked the participants about the snooze function. Despite our clear explanations, some might not have had prior experience or a clear understanding of this function, which could influence their comparisons with the pinning feature. Fifth, our study was conducted during the school break and the participants were mostly young Android users in authors' home country. Therefore, for students the types and experiences of notifications might differ from regular days. Similarly, our results could vary when compared to those from iOS users, older individuals, and people from different locations.

8 CONCLUSION

In this paper, inspired by existing literature advocating various notification management features, we conducted an in-depth study to evaluate the effectiveness of three such features: pinning, categorizing, and sorting. Our investigation reveals the actual usage of these features by participants and their qualitative assessments and feedback. Our findings reveal that pinning emerged as the most frequently used feature among the three features. Moreover, the study uncovers nuanced preferences between manual and automatic modes, contingent upon the specific feature in question. While manual sorting was rarely employed, participants frequently relied on automatically generated categories for navigating their notifications. However, they did not find automatically sorted sequence of notifications to be superior to the original sequence. These insights suggest that user preference

for manual versus automatic modes is not uniform, but rather is influenced by the particular feature being considered. Additionally, our results illuminate the multifaceted utility of pinning, extending its role beyond mere deferral of notifications to also include functions such as preventing accidental notification removal and facilitating future information retrieval. Therefore, this study contributes valuable empirical evidence to the notification research, offering guidelines for the design of more effective notification management systems.

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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–27.

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

[5] 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.

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

[7] 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*. 9–10.

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

[9] 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

[10] 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

[11] 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*. 1–18.

[12] 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*. 1–15.

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

[14] 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

[15] Mihaly Csikszentmihalyi, 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.

[16] 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

[17] 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.

[18] 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

[19] 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

[20] 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

[21] Danyel Fisher, AJ Brush, Eric Gleave, and Marc A Smith. 2006. Revisiting Whittaker & Sidner's" email overload" ten years later. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. 309–312.

[22] 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

[23] Catherine Grevet, David Choi, Debra Kumar, and Eric Gilbert. 2014. Overload is overloaded: email in the age of Gmail. In *Proceedings of the sigchi conference on human factors in computing systems*. 793–802.

[24] 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*. 1–21.

[25] 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*. 93–102.

[26] 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*. 27–30.

[27] 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*. 1–16.

[28] 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

[29] 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.

[30] 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

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

[32] 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.

[33] 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

[34] Tianshi Li, Julia Katherine Haines, Miguel Flores Ruiz de Eguino, Jason I. Hong, and Jeffrey Nichols. 2021. Alert Now or Never: Understanding and Predicting Notification Preferences of Smartphone Users. *ACM Transactions on Computer-Human Interaction* (2021). https://doi.org/10.1145/3478868 Just Accepted.

[35] 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*. 1–13.

[36] 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

[37] 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

[38] 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

[39] 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

[40] 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

[41] 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

[42] 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

[43] Martin Pielot, Bruno Cardoso, Kleomenis Katevas, Joan Serrà, Aleksandar Matic, and Nuria Oliver. 2017. Beyond interruptibility: Predicting opportune moments to engage mobile phone users. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–25.

[44] 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

[45] 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

[46] 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

[47] 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.

[48] 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

[49] 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.

[50] 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*. 2269–2278.

[51] Eran Toch, Hadas Chassidim, and Tali Hatuka. 2020. Can you Turn it Off? The Spatial and Social Context of Mobile Disturbance. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 91:1–91:18. https://doi.org/10.1145/3415162

[52] 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*. 1–15.

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

[54] 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

[55] 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.

[56] 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

[57] 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

[58] 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

[59] 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

[60] Steve Whittaker and Candace Sidner. 1996. Email overload: exploring personal information management of email. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 276–283.

[61] 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

[62] 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*. 5346–5360.