

Reflective Spring Cleaning: Using Personal Informatics to Support Infrequent Notification Personalization

Izabelle Janzen
bjanzen@cs.ubc.ca
University of British Columbia
Vancouver, BC, Canada

Joanna McGrenere
joanna@cs.ubc.ca
University of British Columbia
Vancouver, BC, Canada

ABSTRACT

Distracting mobile notifications are a high-profile problem but previous research suggests notification management tools are under-used because of the barriers users face in relation to the perceived benefits. We posit that users might be more motivated to personalize if they could view contextual data for how personalizations would have impacted their recent notifications. We propose the ‘Reflective Spring Cleaning’ approach to support notification management through infrequent personalization with visualization of collected notification data. To simplify and contextualize key trends in a user’s notifications, we framed these visualizations within a novel who-what-when data abstraction. We evaluated it through a four-week longitudinal study: 21 participants logged their notifications before and after a personalization session that included suggestions for notification management contextualized against visualizations of their recent notifications. A debriefing interview described their new experience after two more weeks of logging. Our approach encouraged users to critically reflect on their notifications, which frequently inspired them to personalize and improved the experience of the majority.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI; Information visualization; Mobile devices.*

KEYWORDS

personalization, visualization, notifications, interruption, personal informatics

ACM Reference Format:

Izabelle Janzen and Joanna McGrenere. 2022. Reflective Spring Cleaning: Using Personal Informatics to Support Infrequent Notification Personalization. In *CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3491102.3517493>

1 INTRODUCTION

Increased distractions through notifications, often in the form of pop-up messages on a mobile phone, is a problem that has been of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-9157-3/22/04...\$15.00
<https://doi.org/10.1145/3491102.3517493>

great research interest and a focus of public discourse [36, 39, 55, 64]. Such distraction has been shown to cause stress to users and can be particularly frustrating if the notifications are not urgent, relevant or useful [4, 16, 24, 47, 54, 66]. However, users have many contextual and individual differences in what notifications they are receptive to receiving despite this stressful distraction. For example, the importance of seeing work messages on the weekends or social messages while working vary for each user [4, 38, 54, 60, 67]. In this work we explore the problem of designing a notification management system where users can actively define which notifications they are receptive to receiving despite the distraction they cause [4, 6, 15, 85].

Prior notification management tools in both industry and academia have used software personalization to support users actively defining their notification preferences [26, 34, 35, 53, 64, 80]. The techniques used are very diverse but usually involve the user describing which types of notifications they do not want to receive, such as by creating rules to filter unnecessary notifications [6, 15, 31, 37, 65, 82, 85]. However, many users avoid personalizing due to barriers such as the time required to create settings until they are convinced the benefits of personalizing are worth the effort [8, 53, 54, 78]. For example, they might be triggered to personalize by receiving a sudden increase in particularly disruptive notifications [8]. These triggering events are often necessary to motivate personalization because problematic habits of technology use commonly become unconscious and normalized [25, 46, 79]. Related, wellness systems using journals or visualizations of usage data have been shown to encourage reflection on user habits and critically challenge user assumptions about the way they use their mobile technology [1, 25, 46, 68, 79].

Deep reflection about notification use could motivate users to personalize their notifications, but reflection is a complex, non-linear process that takes focused time in specific activities, such as through writing a journal or viewing personal informatics [1, 10, 25, 79]. Personal informatics (PI) systems are tools that enable users to track, visualize, and analyze data about personal behaviors like web browsing habits to encourage deeper reflection on how to improve those behaviors [20, 51, 70, 79, 83]. Prior notification personalization tools, to the best of our knowledge, have not explicitly provided opportunities for focused reflection which may limit user motivation to personalize [4, 53, 78, 87]. For example, users may assume notifications are not worth spending the effort to personalize if they underestimate the number or disruptiveness of notifications they receive [8, 53, 54]. While reflection may increase user motivation to personalize notifications, research on PI systems in other domains, such as mental health, has already suggested that conducting thorough PI data collection and deep reflection presents

substantial barriers to users that are not intrinsically motivated by self-tracking [72, 73]. The goal of our work was therefore to explore how a notification management system could provide PI visualizations of notification use that provide opportunities for explicit, useful reflection, while minimizing the complexity and overhead of using those visualization components so that users would not be further deterred.

PI of notifications was very briefly explored in one prior study [88]; they laid the technical groundwork to create visualizations of notification usage data but only did a very small ($n=3$) evaluation of impressions of graphs of their recorded data [88]. To accomplish our goal of minimizing the complexity and overhead of using PI visualizations of notification data, we explored the idea of a simplified data abstraction. Data abstractions are theoretical tools widely used in the visualization literature that map raw data, such as individual notifications, into semantically relevant categories for visualization [58]. Identifying a data abstraction and associated encoding of that data into a visualization that captures the most relevant trends in notification data, while remaining as simple and easy to understand as possible, was the focus of the design process for our visualization components.

To allow participants to actually personalize their notifications, we integrated these reflective visualizations with pseudo artificially intelligent driven suggestions for notification management. This personalization approach, where the user and the system “work together” to achieve a personalization, often referred to as mixed-initiative personalization, is the current state of the art notification management technique [34, 53]. For example, Prefminer is a system that learns how users respond to notifications and suggests different filtering rules to remove irrelevant notifications such as “filter Facebook notifications with ‘candy’ and ‘crush’ in the title” [53]. These suggestion based approaches do reduce the barriers to personalization by leveraging machine intelligence, but users still need to be engaged enough in notification management to fully interact with the suggestions.

Our core research questions are: **RQ1** How would simple visualizations of notification use in a notification management tool influence users’ motivation to personalize their notifications? **RQ2** How would such a system affect users’ notification management practices and experience?

To answer these questions, we created a novel notification management design called Reflective Spring Cleaning. This design approach involves visualizations of the user’s notification usage data, personalized suggestions for notification management, linked highlighting to show how notifications will be impacted by the suggestions, and a prompt to Spring Clean very infrequently, such as every six months. Our design approach thus differs substantially from prior mixed-initiative suggestion systems that frequently, such as one suggestion every few days or weeks, give users individual personalization suggestions with no visualization. For the visualization itself, we created a novel “who-what-when” data abstraction that maps prior domain knowledge of what types of notifications users value to three visualizable facets of notifications. Our data abstraction was created through extensive synthesis of prior qualitative work and was refined using low-fi design iteration [4, 54, 58, 66].

We conducted a four-week longitudinal study to evaluate the effectiveness of our Reflective Spring Cleaning approach and data abstraction where participants passively logged notifications for two weeks, then personalized with Spring Cleaning, followed by two more weeks of passive logging, and finally a debriefing interview where they reflected on their new experience with notifications.

Our findings revealed that the Reflective Spring Cleaning design prompted participants to reflect critically on their habits of notification use, which raised their awareness of their need and ability to personalize those notifications. Further, the facets of our data abstraction (Who, What and When) all provided useful insights about notifications. Reflecting on these abstractions supported utilitarian notification management, satisfied participants’ brief intrinsic curiosity, and helped them to minimize perceived sources of stress from distraction. Finally, individual differences in past notification management behavior influenced participants’ willingness to continue engaging in Reflective Spring Cleaning. The primary contributions of this work are:

- Design of the Reflective Spring Cleaning approach that combines research from mixed-initiative personalization systems [53, 69] with visualizations of PI data [88] to encourage critical reflection on notification use and motivate users to overcome barriers to personalizing notifications [8, 78].
- Evaluation of the Reflective Spring Cleaning design in a structured study involving participants’ own notification data, showing that it supports several motivations for reflection and can improve participants’ experiences with notifications. This improvement in experience was observed despite the potential barrier reflection adds compared to prior notification management systems [53, 69].
- Exploration of the novel data abstraction created in the Reflective Spring Cleaning design to demonstrate that the data abstraction facets provide useful insights that are actionable in combination with the included personalization features. This combination of PI and personalization may have broader implications for making reflective systems more actionable in general [20, 83].

2 RELATED WORK

We provide an overview of research into what notifications users are receptive to receiving, personalization techniques to apply those preferences to a notification system, and ways to support reflection through personal informatics.

2.1 Factors Impacting Receptivity to Notifications

The factors that affect which notifications users are receptive to receiving despite the distraction they cause are complex, contextual and difficult to precisely predict [4, 24, 54, 69, 84]. Work has identified both contextual and content based factors impacting which notifications users are receptive to or otherwise value. The contextual factors usually revolve around what primary tasks the notification is distracting the user from, such as avoiding distraction from work during a family dinner [4, 16, 24, 31, 54, 84]. Content factors are many and varied but can include different senders. For

example, close social contacts are usually more important than acquaintances, strangers or automated systems. Other content factors include urgency to respond, relevance to work or even just interest in the content [4, 24, 54, 84]. Importantly, these contextual and content factors have been shown to be dependent on each other and subject to strong individual differences, necessitating user-level personalization. For example family messages might typically be more important, but in certain contexts, like work, they still might need to be ignored for some users [4, 54].

An early approach to improve the experience of these systems was using a machine to automatically deliver notifications when users are least disrupted by them [3, 24, 36, 76]. Commonly, this involves delivering notifications when the user is already switching between tasks [24, 76]. However, this approach has fallen somewhat out of favor, possibly due to users fearing that they may miss key notifications or that they lack control when they are not in the loop [4, 41]. Research to understand which notifications should be deferred in these systems often derives meaning about which notifications need to be delivered promptly based on in-the-moment experience sampling methodologies or behavior analysis [40, 53, 54, 54, 66, 81]. However, users may decide to further personalize their notifications manually and might have difficulty remembering which notifications were problematic days or weeks after receiving them. Our work explores to what extent deeper data-driven reflection during notification management might encourage users to personalize, such as by helping them remember which notifications were problematic.

2.2 Approaches to Software Personalization and Notification Management

Software personalization involves a system adapting to the user needs and differences, such as which notifications they are receptive to receiving, and has been explored in many domains [8, 12–14, 27, 28, 48]. Manual personalization has been the most common prior approach to software personalization. However, users experience barriers to personalization that triggering events, such as particularly frustrating errors, push users to overcome despite the effort required [8]. The barriers are sufficiently high that few users tend to personalize software [8, 53, 78, 87]. However, users still report substantial frustration with notifications, thus pushing the research community to try to reduce these barriers [4, 66, 78, 87]. Typically, machine intelligence is used to learn useful adaptations to the existing settings based on what the system knows about each user [8, 12–14, 27, 28, 48]. Research has identified a spectrum of techniques for applying this learning from fully adaptive systems (where the machine automatically optimizes the settings) to mixed-initiative systems (where the machine suggests improvements to the user) that have all been applied to notifications [8, 12–14, 26–28, 35, 48, 64, 80, 86]. These come with tough design trade-offs such as a lack of control in adaptive deferment systems, or increased distraction in mixed-initiative [6, 15, 24, 53, 86].

While there are a variety of personalization approaches to which we could have explored adding focused reflection, mixed-initiative suggestions are the state of the art personalization technique for notifications [4, 34, 41, 54]. Further, mixed-initiative suggestions reduce the barriers to personalize notifications, provide greater user

control than automated deferment systems, and the suggestions themselves could serve as a small prompt to reflect more deeply on notification use [6, 8, 13, 53]. For example, in industry, iOS15 is currently in developer beta and explores mixed-initiative personalization through a new “focus” mode where on-board artificial intelligence will automatically determine the current phone context such as working or commuting [34]. Users can choose and then customize a suggested “focus” such as ignoring non-work messages while working, similar to prior do-not-disturb features on both Android and iOS. Focus mode will likely improve a user’s notification experience, though users will still need to be engaged in notification management enough to properly customize which foci should be applied in which circumstances [24, 27, 53, 82]. Our work explores to what extent users’ decisions about how to deliver notifications in a similar system would be influenced by reflection on their notification use, and more specifically, whether reflection would encourage users to engage with the system.

2.3 Challenges in Reflective Design and Personal Informatics

Reflection involves serious thought and consideration for purposes such as learning or critical review and needs to be done in appropriate environments to achieve differing levels of depth and nuance [25, 79]. Particularly deep reflection that involves identifying and questioning deeply held assumptions is described as critical reflection [56]. For example, mental health and wellness researchers have argued that users should be encouraged to deeply reflect on how healthy and necessary their use of mobile technology is [1, 9, 23, 46, 79]. A wide variety of tools have been developed to support deep reflection on health and wellness by monitoring concrete data, such as through guided journaling [10, 68], using personal informatics to track personal behaviors and mood [18, 51, 77, 83, 88], or visualizing data about peer technology use [22, 46, 59, 71]. Research suggests that deep reflection is either a primary outcome of research [7, 10, 23, 25, 52, 75] or a way to encourage the user to refine and improve their personal behavior [20, 51, 83]. For example, research has shown that reflection on personal informatics can help raise consciousness of potential issues and maintain new behavior through self-monitoring, which are key steps in established models of behavior change [20, 83].

Prior research has identified a number of key research challenges in prior reflective interfaces. Without an intrinsic motivation to self-track, such as strong curiosity about one’s quantified self or a need to manage a chronic health condition, general users may quickly stop tracking for many reasons [19, 33, 72, 73]. For example, onerous tracking and analysis requirements, a lack of actionable takeaways, or difficulty identifying what data should be tracked may cause frustration and dropout [19, 20, 62, 63, 83]. While some users may feel satisfied in what they learned and be happy to disengage from self-tracking, others do so out of frustration [21, 42, 43]. To minimize the chances of this occurring, research has explored ways to simplify many elements of the self-tracking process, such as by displaying subsets of recorded data and using simplified visual representations [19, 29, 74]. We sought to extend this line of research to the domain of notifications, specifically by creating a simplified data abstraction

that would map notification data into visualizable categories that summarize the most important trends within notification use.

While we were not specifically aiming to improve prior personal informatics systems, we did worry that many users of notification management software may not have extensive experience with other self-tracking tools. Supporting users who are not already experienced with and intrinsically motivated by self-tracking has been a key and on-going research challenge within personal informatics [23, 42, 72]. We wondered to what extent our approach of simple self-tracking elements included in a notification management system would generate deep reflection for users motivated more by notification management than self-tracking in general.

3 DESIGN PROCESS

The design process to create our prototype followed two main stages. In the first stage, we used a pre-study survey to generate notification management suggestions that our mixed-initiative design would present to users. In the second stage, we used iterative prototyping techniques to explore what potential data abstractions could encourage users to reflect on notification use and how those reflective components could be included alongside our suggestions.

3.1 Pre-Study Survey to Generate a Corpus of Useful Suggestions

In order to evaluate whether PI visualizations encourage participants to engage in notification management, we needed a corpus of reasonable suggestions that participants could react to while reflecting. We had tentatively created a list of 20 potentially useful notification management suggestions based on our analysis of prior work on qualitative and quantitative trends in notification receptivity, as well as our own personal experience [4, 41, 54]. For example, prior research has shown some users choose to hide their devices to avoid being distracted while working, so this list included silencing personal notifications in work hours [4, 41, 54]. Other, more contextual suggestions were based on research about what notifications users value, such as important work messages or those from close family members that we suggested to highlight with a special sound [4, 41, 54]. Variations of these and other common notification preferences were explored and varied, such as by trying more specific versions that applied to different contexts. The full list of pre-study survey suggestions is shown in Appendix A of supplementary materials. To ensure that our longitudinal study participants would not be distracted from the visualizations by obviously poor suggestions, we conducted a short pre-study survey to identify which of the 20 suggestions participants would find most useful.

In total, this pre-study survey was distributed to 140 participants (mean age=24.9, SD=5.7; 73 students along with various forms of knowledge workers such as clerks, web developers, analysts or consultants) through convenience sampling on online recruiting forums and the university paid studies board asking “For the following list, think about whether or not you would accept such a suggestion.” Users could select “yes,” “no,” or “not sure” for each suggestion shown in Appendix A of supplementary materials. Ethical approval for the study was given by our university research ethics board. Seven suggestions were candidates for inclusion in the

longitudinal study because they met the following criteria: They could be mimicked with current phone settings; a minimum of 50% of the pre-study survey participants would accept them; and they changed notification delivery modality rather than fully removing notifications. The final criterion essentially avoided fully filtering out important messages, and was deliberately a conservative approach. Refined versions of these seven suggestions are shown in Table 1 and were displayed to users in the longitudinal study presented in Section 4.

3.2 Iterative Refinement and Final Design of the Reflective Spring Cleaning Design

Our design process to create the Reflective Spring Cleaning concept was highly iterative through brainstorming, low-fi sketching, and wire framing. For example, we created a variety of sketches of possible categorizations of notifications based on prior work about which notifications users are receptive to, or willing to receive despite the distraction they cause. One such sketch is shown at the top of Figure 1. We used simple visual encodings in each sketch, such as bar charts or timelines, to reduce the barriers to conducting reflection and personalization [8, 58]. For example, we explored breaking down notifications in a bar chart grouped by how socially close the sender was to the recipient, or which context the notification was received in, because both factors strongly impact users’ willingness to receive those notifications [4, 54]. Further, we explored a large number of possible labels to communicate the often fuzzy boundaries between different groups of people sending notifications. We conducted informal cognitive walkthroughs where we role-played a user personalizing their notifications for the first time and viewed each sketch describing their notifications. For each sketch, we discussed what major takeaways a user would receive from viewing the graphs and whether those insights could impact the decision to accept the seven suggestions selected for use in the longitudinal study from our pre-study corpus [4, 41, 54]. The key high-level design insight from these walkthroughs was that it was difficult to understand trends in a participant’s notifications if they were categorized on more than one factor affecting notification receptivity all at once. For example, prior work suggested differentiating urgent “Mission Critical” notifications from non-urgent “Nagging” notifications, but urgency reflects both variance in who sent the notification and in what context it was received [4, 41]. We explored various approaches where these factors were split into several complementary but independent data abstraction facets which became the basis for our “who-what-when” data abstraction. Examples are shown in Figure 1. The potential cognitive cost of viewing several graphs of notification data was a key reason our design approach settled on very infrequent personalization in the Spring Cleaning concept.

To explore the value of the Reflective Spring Cleaning concept and our data abstraction, we built a semi-functional HTML prototype leveraging pre-existing tools [86]. This allowed us to focus predominantly on the design of the prototype, rather than technical implementation of notification interception and logging which has been well explored in prior research [53, 54, 86, 88]. We built on prior tools by processing the raw notification data from participants, logged using open sourced tools to roughly classify it into

Table 1: The 7 notification management suggestions shown to participants in the longitudinal study

Suggestion Text	Show to participants in longitudinal study if
Highlight messages from priority contacts	>4 notifications per day from priority contacts
Silence personal notifications during working hours	>10% of notifications in work hours are personal content
Highlight work notifications	<40% of notifications are from work
Silence system notifications	>3% of total notifications are system notifications
Silence social media notifications	>4 social media notifications per day with no sender
Silence work notifications during off-work hours	>10% of notifications in off-work hours are from work
Highlight work notifications during off-work hours	<10% of notifications in off-work hours are from work

the categories visualized in our data abstraction facets [86]. For example, the data did not include any predefined sender category and many messaging apps differed in how this data was reported, such as in the ‘title’ attribute or just as part of the text body. Our prototype includes a rough estimate of time spent reading notifications based on prior work and a short notification categorization form (around five minutes to complete) for participants to identify work notifications, priority contacts and primary working hours.

We piloted our longitudinal study (Section 4) with seven participants in order to assess the seven suggestions from the pre-study survey. The pilot participants viewed all seven particularly useful suggestions in the pre-study survey and we asked them to describe which were most relevant to their needs and why. We turned these insights into relevance guidelines for our program to know when a suggestion might be useful. For example, we might suggest silencing personal notifications if a substantial portion of their notifications in work hours are personal according to a rough threshold, such as >10%. The final prototype would then show three to four relevant suggestions to each user in a random order based on the guidelines in Table 1; at least three suggestions were presented to each user. The final prototype is shown in Figure 3. The final data abstraction facets we showed to participants are as given in Figure 2 and below.

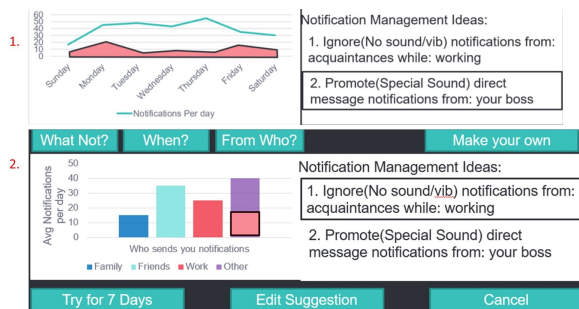


Figure 1: Examples of early design sketches for Reflective Spring Cleaning. 1. Shows an early view of when notifications are received. 2. A later more refined sketch that splits apart the emerging abstractions, currently showing the From “Who” facet.

3.2.1 What. This facet describes what types of apps sent the participant notifications and is visualized with a bar chart of the categories. Categorizations of apps have been made in both industry and academia [54, 78]. For example, the Google Play Store lists 26

categories of apps to download [30]. Through our informal cognitive walkthroughs we explored what insights participants might gain from reflecting on these commercial breakdowns of apps, but many of the categories were assigned for apps that rarely send notifications. Our final design was inspired by the types of apps that Pielot reported sent the most notifications in their large-scale observational studies of notifications from over 40,000 users, as it remains the most comprehensive and inclusive study of notification usage [66, 78]. The “What” facet includes 11 categories (System, Social media, Games, Email, Text, Voice, Calendar, News, Tools, Market and Other), as shown in Appendix A.

3.2.2 Who. This facet describes the types of people who sent the participant notifications based on prior work exploring notification receptivity and social relationships [4, 54, 60]. Social relationships are often complex and individual, so we prompted participants to describe which contacts they valued most [4, 54, 60]. The final categories are: A group of important personal contacts provided by participants called “Priority Contacts.” “Work Contacts” are identified by notifications participants receive from a work messaging app, email address, or work contacts such as a boss. Remaining notifications that identify a sender called “Other Contacts”. All other notifications are called “No Sender Given.”

3.2.3 When. This facet describes the notifications participants received while working or not, which strongly impacts notification receptivity [4, 24, 54, 66]. Our “When” categorization split notifications into two contexts divided by two types of notifications. The contexts are: “Primary Work Hours”, as defined by the participant for each day of the week through a form completed just before Spring Cleaning, and its complement context, “Off-work hours.” These contexts are further broken down by two types of notification content: “Work Content” and “Personal Content.”

Lastly, we enhanced the visualizations of these facets by allowing participants to compare the different weeks of data as shown in Figure 3. Each visualization allows participants to click on the categories to refine the data shown. For example, if users clicked on the “Other Contacts” category in the “Who” facet they saw a bar chart of all the senders that were not work or priority.



Figure 2: The three facets of our data abstraction. Red bars are all 0 because this user is not currently comparing different weeks of data.

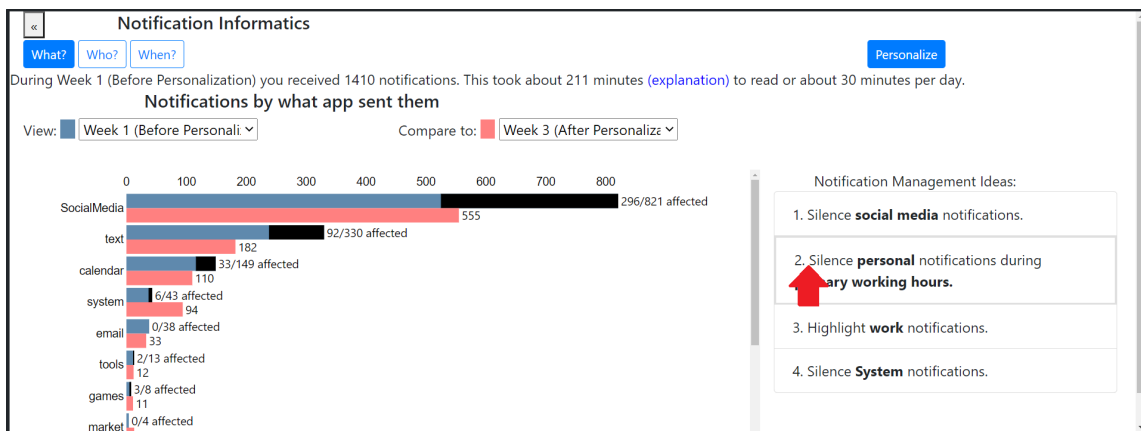


Figure 3: The final Reflective Spring Cleaning Interface showing real anonymized data. Here the user is viewing the “What” facet, comparing Week 1 to Week 3 and has clicked the “silence personal notifications during working hours” suggestion indicated by the red arrow. The black bars indicate the notifications that would have been impacted by the clicked suggestion had the personalization already been in place; the red bars show what notifications were received in Week 3 after applying the suggestion.

4 METHODS FOR THE LONGITUDINAL STUDY OF REFLECTIVE SPRING CLEANING

To understand the usefulness of and motivating our approach of including PI visualization in notification personalization, we evaluated our Reflective Spring Cleaning prototype in a four-week longitudinal study.

4.1 Participants & Recruitment

Participants were recruited through online calls for participation on classified forums and university paid studies boards using convenience sampling. Ethical approval for the study was given by our university research ethics board. Technological restrictions

meant participants needed to use Android devices and be able to identify an email account, messaging app or set of contacts primarily devoted to work so we could properly process their data. Several Android phone brands including Xiaomi and Huawei were excluded because of known issues with the logging app that made data gathering inconsistent. We also required participants to be 18+ years of age, speak English and “feel they receive a lot of notifications and might be interested in better managing them.” This requirement targeted smartphone users who could be open to the idea of notification personalization and are the key demographic for notification management.

Our final dataset included data from 21 (mean age= 26.5, SD=8.5) participants, nine of whom completed the optional email follow up. One additional participant dropped out of the study due to their concerns that the suggestion they wanted to apply, silencing non-work notifications, could interrupt their small business if implemented with current settings. They are not included in our dataset. As has been observed in prior work, recruitment from university paid studies boards run by the psychology department can be gender biased in favour of women, which was the case in this study with four men and 18 women participating [61].

Our recruitment criteria seeking participants who were somewhat dissatisfied with their notifications led to our participants being relatively engaged in notification management. 13 participants reported at least occasionally personalizing notifications through the phone's notification settings, do-not-disturb mode, or a stay focused app. Of the remaining nine participants, seven described occasionally uninstalling apps that send particularly annoying notifications such as games they no longer play, while the remaining two did not manage notifications. While most participants described having mildly negative experiences with their notifications, several participants felt overwhelmed and kept their devices permanently on silent. While we did not have explicit recruitment quotas, we did make a conscious effort to recruit a roughly even split of participants who are students and professionals (11 students, 10 professionals) to evaluate our design with a broad set of users. Of the professionals, their occupations varied and included: mechanic, research coordinator, copywriter, admissions officer, admin assistant, teacher, and territory manager.

4.2 Procedure

An overview of the procedure is shown in Figure 4. Step 1. Two Weeks of Passive Logging: After confirming eligibility for participation via email, participants were instructed to install the notification logging app [88]. Participants were then instructed to use their phone normally while the system passively collected data on their notification use. Passive data collection lasted two weeks and we checked data was being recorded properly the day after installation and half-way through the period.

Step 2. Spring Cleaning Session: Participants were asked to manually export the log data from their devices. The next day participants participated in the Spring Cleaning session via Zoom, together with a research member. After using our notification categorization form to provide a work email, priority contacts, and working hours, participants were shown the Reflective Spring Cleaning prototype via the researcher's shared screen, which they could interact with remotely. To encourage initial reflection, the suggestions for notification management were hidden to start and participants were instructed to explore and talk-aloud anything that was interesting or surprising in their data. Initial reflection took between three and seven minutes (avg 4 min 51 sec). Participants then completed a brief semi-structured interview on the things they learned about their notification use and their impressions of the data abstraction facets. Afterwards, participants were shown the notification management suggestions and asked if they wanted to apply any of them. If so, the researcher guided them through the instructions to setup that suggestion on their device. Participants were also asked if they

had any other changes they wanted to make to their notification settings and to explain why they wanted to change those settings. Participants who did think of such changes were asked to make them during or after the session. If asked by the participant, the interviewer described where to find the settings they wanted.

Step 3. Two More Weeks of Passive Logging: Participants were instructed to use their phone normally for the two weeks of passive logging after the first reflection session. Participants were instructed to feel free to make any further changes they wanted to their notification settings, but to keep track of any changes made.

Step 4. Debriefing Interview Involving Reflection: Participants' notification data was exported again. Participants completed a final semi-structured reflection session via Zoom. Participants described their impressions of how their notification experience differed since personalizing. Participants then explored all the same visualizations as before with updated data. Participants were informed they could use the drop down bars at the top of the visualizations to compare different weeks of data, but were not required to do so. Figure 3 shows a participant comparing their data from before Spring Cleaning (blue) to after Spring Cleaning (red) on the what abstraction. The black highlights show the impact the indicated suggestion was predicted to have before personalizing. After reflecting, participants discussed the design as a whole, the effectiveness of the personalizations they applied, and whether they might keep using them.

Step 5. Brief Optional Email Follow-up: A month after the final reflection session we sent participants a two question optional follow-up by email. Participants described whether they continued to use the personalizations they applied and whether they had made any further changes to notification settings.

Video recordings of the Spring Cleaning session (Step 2) and the final debriefing interview (Step 4) were transcribed verbatim. We also recorded demographics data and asked participants to self-rate expertise and usage of various devices including phones and tablets in Step 2.

4.3 Data Analysis

Thematic analysis of transcript data progressed iteratively in two main phases, under the framework by Braun and Clarke [11]. Our first phase of analysis was more targeted and focused on understanding what sorts of insights participants were gaining about their notification use while reflecting on the data abstraction. Understanding what insights were generated was key, as the usefulness of Reflective Spring Cleaning was dependent on whether generated insights were useful for personalization. In this phase, the lead researcher would manually scrub through transcripts looking for times participants talked aloud during the reflection session or described learning something about their notifications in comments. These potential insights were then categorized by whatever design element may have generated them, and qualitative codes were applied to describe what type of comment or insight they were. For example, several insights were described as "surprise over who sends the most messages." Special focus was given to categorizing any comment where participants either explicitly or implicitly implied their notification personalization decisions had

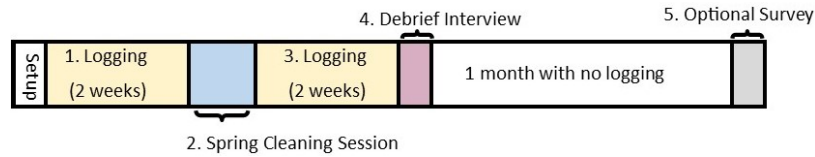


Figure 4: Overview of the procedure for the longitudinal reflection study.

changed. We then analyzed these codes into candidate themes of insights gained.

A second phase of thematic analysis was then conducted using inductive open coding. We focused on a descriptive analysis of the overall impact of the design, concerns participants had over notifications, and how their experience with notifications shifted over the course of the study. The primary consideration for reliability was increasing the nuance of the analysis through multiple perspectives by having input at various stages from as many researchers as possible, all of whom could use their own unique experiences and perspectives on how they use and manage notifications. The goal of this iterative analysis was to incorporate these multiple perspectives to approach crystallization [11, 17]. In total, the lead researcher coded every transcript, two other researchers independently coded three transcripts each, and a final researcher coded a further two transcripts. Candidate themes were then developed and refined through discussion in a series of meetings of the research group, along with several other HCI faculty in associated institutions.

5 FINDINGS

We first provide an overview of how participants personalized with the Spring Cleaning tool and its impact on their notification experience. We then describe through three qualitative themes the impact of reflection on notification management and participants' motivation to manage notifications. Theme 1 describes the reflection engaged in by participants. Theme 2 describes their motivations to conduct this reflection. Theme 3 describes their willingness to engage in further notification management after Spring Cleaning.

5.1 Overall Usage of Spring Cleaning Tool

We give an overview of how the Spring Cleaning tool was used by participants. Further, we describe how these changes influenced participants' experiences with notifications and improved that experience for the majority.

5.1.1 All participants conducted personalization, 9 accepting our suggestions and 12 creating their own personalizations. In the personalization session 9/21 participants applied one of our suggestions to their phone (four highlighted work notifications with a special sound, five silenced work or personal notifications at suggested times). The rate of acceptance of our suggestions was lower in this study than in the pre-study survey where, for example, 75% of participants wanted to silence non-work notifications during working hours. Some of this difference is likely because the pre-study survey did not require participants to actually try the personalization and risk possible errors.

All participants who did not apply one of our suggestions (12/21) applied other personalizations, usually based on insights they gained while reflecting. Of these 12 participants, six reported that the "What" facet showed too many notifications coming from some sources. They described uninstalling apps or notification services such as games or social media from those sources in the personalization session. Two participants reported the "Who" facet showed too many notifications from certain types of group chats, and silenced them or set them to only notify the participant when tagged. P3 - "In the past I didn't silence [group chats], because they're not that bothersome... But then seeing them visualized... they take up a lot of space and I should definitely change the settings." The remaining four participants reported their reflection prompted them to do a broader cleanup of their notifications over the next few days. P15 - "I blocked all the notifications that weren't useful for me. Social media, some of the emails and all the other stuff that I could disable I disabled." No participant who accepted one of our suggestions applied other personalizations.

5.1.2 The majority felt the personalization conducted improved their experience. The notification management suggestions we provided participants changed the modality of how notifications were delivered but did not change the total number of notifications delivered, so we did not anticipate significant quantitative differences and none were observed as shown in Figure 5. Descriptively, participants received fewer notifications in the two weeks after Spring Cleaning (avg 3669) compared to the two weeks before Spring Cleaning (avg 3989). However, natural variation in the data set was high (SD 2451), such as work notifications shifting during holidays or around deadlines. P9 - "I'm not working right now, just finishing up exams... So I haven't been receiving [as many] messages related to school or work anyways."

It was encouraging to see that more than half the participants (13/21) reported a subjective improvement in their notification experience after Spring Cleaning and that this group was evenly split between those accepting our suggestions (6) and those applying their own personalizations (7). Of these 13 participants who reported improved notification experience, four described being less distracted and more productive, three said they were less overwhelmed and stressed by notifications, two described having better boundaries between work and personal time, and one described receiving fewer notifications not relevant to their current tasks. The remaining three participants simply described that their notifications just felt better than normal. We estimate that the suggestions participants accepted impacted similar proportions of their notifications (avg 1644 or 31% of their notifications) to the personalizations participants created for themselves (avg 781 or 25% of their notifications). Overall, participants personalized an average of 28% of

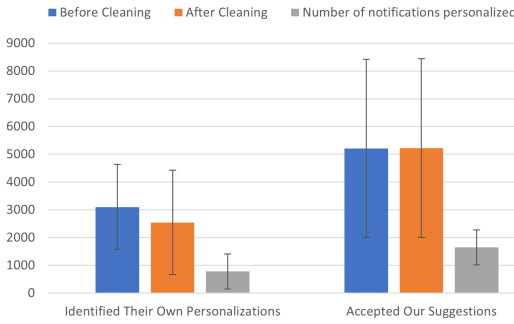


Figure 5: Average number of notifications received by participants broken down by whether they accepted our suggestions or created their own personalizations, with bars for standard deviation. Estimated average number of notifications impacted by those personalizations is visualized in grey.

their notifications, which is comparable to prior mixed-initiative systems. For example, Prefminer reported filtering an average of 16% (12/71) of participants’ daily notifications [53].

While typically participants considered the overall improvements from Spring Cleaning as small positive refinements, in some cases, this improvement was considered very positive if the personalizations impacted a large and noticeable number of notifications. In the most dramatic example, P22 - “cancelled every possible notification I could” and observed a drop from 4335 notifications in the two weeks before using Spring Cleaning to 1545 in the two weeks after. We roughly estimate 2250 notifications of the total 2790 reduction can be attributed to personalization. For example, several apps were uninstalled which removed all their notifications. While debriefing, P22 expressed “I’m blown away by how much just a couple simple changes in those notifications can improve productivity and lighten the load a bit.” Of those participants who did not report improvement, they usually described the changes as being easy to overlook. For example, changing the ringtone for work notifications wouldn’t matter if they always checked their phone when it buzzed. Some participants also described that the changes were not noticeable due to natural variation in their number of notifications. For example, P16 described turning off system updates and weather reports they didn’t need, which removed 1060 notifications, but their overall number of total notifications increased (6183 before, 6782 after) due to greater messaging. P16 - “I was getting less from certain things but in terms of overall effect, it didn’t seem very different.”

5.1.3 Summary. All participants personalized their notifications with Reflective Spring Cleaning either by accepting our suggestions or applying other personalizations inspired by their reflection. The majority felt their personalizations improved their notification experience.

5.2 Critical reflection about notification usage was prompted

Participants frequently described observing surprising events or trends of notification use while interacting with the visualizations

in the Spring Cleaning tool. These insights prompted participants to critically reflect on how disruptive notifications were in their everyday lives, especially by identifying that they were receiving more unnecessary notifications than they expected. Further, participants linked their enhanced understanding of the cost of notifications to their motivation to apply personalizations in the study. Reflection also influenced which personalizations they thought would be effective to reduce the cost (RQ1).

5.2.1 High volume of unnecessary notifications to remove, often derived from the “What” facet, surprised participants. Most (18/21) participants felt they had underestimated the number of notifications they received, especially related to apps participants felt they didn’t need. For example, P22 noticed they received over 200 notifications from games in two weeks and remarked: “Now the sad part is it makes sense. I receive them multiple times a day. I would love for them to be off my phone during the day.” Participants frequently made these observations about receiving a surprising amount of unnecessary notifications while viewing the “What” facet (15/18). P7 - “The first thing I noticed is that I didn’t think I received this many text notifications... I’ll mute most of them.” In addition to underestimating the overall number of unnecessary notifications, participants also frequently reported realizing through the “What” facet that the largest sources of distracting notifications were different than they expected. P5 - “I thought with the amount of time I’ve been on social media I would have way more [social media notifications]...it’s kind of sad that I have so many system notifications.” Removing these unnecessary notifications was a key goal of many personalizations participants created.

5.2.2 Habits of notification use that contextualized insights were identified, often through the “Who” and “When” facets. Participants described getting a better understanding of the context of their own day-to-day habits of phone use through reflection on the “Who” and “When” facets. For example, P3 remarked while viewing the refined timeline of their notifications in the “When” facet: “I guess Saturday is a day off. It would be kind of expected [to have more notifications] but I didn’t realize it was that big of a spike!” Further, the “Who” and “When” facets provided insights that tended to contextualize problems identified in “What” within a participant’s real-life habits. For example, potentially problematic trends, like seemingly excessive messaging in work hours might actually make sense if it was done in an appropriate context. P10 - “Who’s that [sending me many texts]? Oh, that’s the person we give the blood to. It’s a work thing. Okay, so this isn’t horrible. I am pleasantly surprised because I thought it could be much worse.”

5.2.3 Increased awareness of ability and need to remove unnecessary distraction. Participants often described how prior to the study they had a vague sense that notifications were distracting but were more of a necessary nuisance. P5 - “In the past, I was not managing my notifications because I thought I didn’t let it get to me.” However, after participating in the study participants described being more aware of the costs of receiving many notifications and that they had agency to remove unnecessary notifications rather than let them continue. P3 - “I’m more aware of what customization is available to me, and that it only takes a little bit of effort to make quite a large difference.” Participants being more aware of how to personalize is

perhaps not surprising, given our prototype's limitations and that we instructed participants on how to mimic their desired personalizations using current phone settings. However, we note it here because motivating users to exercise such agency was a key goal of this work and similar personalization training tutorials could be included in future Spring Cleaning systems. Further, some participants reported that part of this agency was due to increased awareness of the problem from reflection rather than just being more trained on how to personalize. P23 - "I guess I am aware of the things that I was seeing that I wasn't aware I was seeing... how actually I had control of that so I didn't have to look at it if I thought it would be not very useful."

5.2.4 Reflection helped participants understand which personalizations would likely reduce distraction. We saw several cases where participants changed their mind about whether to accept suggestions while reflecting on their data. For example, P9, a student, was initially hesitant about accepting our suggestion to silence work notifications during off-work hours but did so after remarking: "Looking at how many [work] notifications I was getting... it's not very many and they're not very urgent... it's usually just [group projects] deciding when to zoom." A few participants also reported how the ability to highlight which notifications would have been impacted by those suggestions on the visualizations changed their decisions on whether to accept them. P2 - "When you get personalization suggestions [like that] you don't take them seriously. When you have it presented to you in a display, it convinces you that you do get a lot. [You're] more incentivised to change your settings."

5.2.5 Summary. Reflection guided by our data abstractions prompted participants to critically examine their notification usage habits. Participants identified that they were receiving more unnecessary notifications than they expected and described having an increased awareness of their ability and need to personalize notifications.

5.3 Individual differences in motivations to reflect

Most participants described having a core utilitarian motivation to reflect in order to better manage notifications. However, individual participants further described having specific secondary motivations such as brief intrinsic curiosity and a desire to minimize stress from distraction.

5.3.1 Identifying accumulated issues provided utilitarian value. The vast majority (20/21) of participants reflected in the final debriefing interview that they saw utilitarian value in Spring Cleaning and would likely use something similar if it was released commercially or as part of the next OS update. For example, P7 said "It would be very useful, especially to know how many notifications [you received] and how distracting those notifications can be." At first glance the magnitude of this finding seemed possibly inconsistent with our overview that suggested only the slight majority of participants reported improved notification experience. What we heard from our participants helped to clarify that the reflection itself was valuable for utilitarian reasons. Many participants described how better understanding their often hard to track notification habits helped them make informed choices about how they managed notifications. P3 - "It's hard to grasp what notifications you're getting unless you

see it visualized like this. Otherwise, when you get a new phone you have to go in and customize everything, it's a lot of work." A few participants also described how after personalizing, later reflection might be useful to monitor ongoing distraction. P20 - "I would use it to make sure I'm on track with my notifications and that I'm not getting unnecessary ones."

5.3.2 Curiosity was a brief intrinsic motivation to reflect. While we did not explicitly ask participants if they would use the design without the suggestions, many participants described how they wanted to infrequently use the visualizations independent of notification management because they were simply curious about how they used notifications and were interested to know more. P10 - "I'm just a data nerd, I like seeing it laid out like this. When you get 20 notifications a day it doesn't seem that overwhelming but then looking at 516 in a week it seems a bit crazy." Meanwhile P1 described how they would view such visualizations on their own because it was interesting to know where they were spending their time: "I definitely find it interesting...It's really hard to realize [where you're spending your time], if you don't track your notifications." Further, some participants described wanting to Spring Clean again some time months into the future to understand how their needs and habits shift over time. P23 - "People's schedules and their preferences change... I would say that I would [Spring Clean] maybe every two months."

5.3.3 Notifications were a source of stress to minimize. Many participants described how they perceived notifications as a frequent source of stress or anxiety in their lives. P3 - "Email notifications are usually something important. [They] give me a shock of anxiety because I'm like 'oh, my God, I have to check what is going on.'" Some participants described how a key outcome of reflection was quantifying the extent of this stress. P13 - "It's a bit stressful now that I see it. That's a lot more [notifications] than I thought [I received]. Beforehand, it didn't stress me out." Further, some participants described wanting to minimize their number of unnecessary notifications to reduce this feeling of stress. P21 - "It starts to feel like anxiety when you open your phone and see a ton of notifications that you don't need."

5.3.4 Summary. Participants displayed several different motivations for reflecting on their notification use with the Spring Cleaning design. Identifying accumulated issues provided utilitarian value while Spring Cleaning, curiosity lead to brief intrinsic interest to understand participants' notification habits independent from personalization, and reflection helped participants quantify and minimize the stress of receiving notifications.

5.4 Degree to which needs were met impacted continued engagement

Participants described how after becoming more motivated to manage notifications by Reflective Spring Cleaning, they usually felt content with how well their refined preferences met their needs. Other participants who actively managed notifications prior to the study reported making further changes to maintain their new experience, but had little desire to reflect further in the short term.

5.4.1 Participants described contentment with refined preferences. After experiencing their refined notification preferences for two weeks, relatively few (5/21) participants felt there were any further changes they wanted to make to their notifications. P2 - *"I can't think of any modifications that I would make as of now."* Most participants described being generally content with how well their notification preferences were met until something substantial changed in their notification use. P4 - *"Maybe once I'm done with college I might change my notification settings again. For now, I don't have any plans."* Further, this sense of contentment with their current notification management practices appeared to persist amongst the nine participants who completed our follow-up survey one month after the final interview. P22 - *"YES [I have continued using my new settings] - I love having less notifications - it's awesome!"* Of the nine participants who responded to the follow-up, five chose not to make further changes. P21 - *"I have not done anything to change my notifications because I feel quite content with them."*

5.4.2 Participants actively managing notifications described maintaining refined experience but limited interest in more frequent reflection. Some participants who actively managed notifications prior to Spring Cleaning described more complex patterns of notification management afterward. While they often described being content with how well their major concerns were addressed by their personalizations, they still considered conducting further notification management. For example, in the study P14 personalized her phone to ignore work IMs in the evening, but was still considering whether she should do the same for work emails. P14 - *"I'm still debating if I should hold on to the emails. I haven't made a decision about that."* Of the nine participants who responded to the followup survey, five had actively managed notifications before Spring Cleaning and four of those five continued making more changes afterwards. Two described updating the preferences of new apps they installed and two made more substantial changes to address further needs. P11 - *"I muted most social media apps because I found they became too distracting in day to day life. Exception for some close contacts."* However, even participants who continued to actively manage notifications described how they felt that reflection on notification use had diminishing returns. P10 - *"I don't know if there's enough detail in notification management to fine tune things to a really meaningful degree every six months, at least for me...I'm consistent in my social media use."* Diminishing returns on the utility of reflection led even the most engaged users who actively managed notifications before and after Spring Cleaning to suggest they didn't need to reflect further for at least a few more months. P15 - *"I would rather do it on a three month basis...Six months, there would be too much data."*

5.4.3 Summary. Individual differences in participants' prior notification management practices influenced the degree to which participants notification management needs were met by Spring Cleaning and their notification management practices after Spring Cleaning. While most participants described feeling content with the new state of their notification preferences, some participants who actively managed notifications in the past continued managing notifications to maintain and refine their experience. All participants expressed limited interest in more frequent reflection.

6 DISCUSSION

We focus our discussion on analyzing the effects of combining simplified personal informatics with notification management and the potential implications this approach might have for future reflective design or personalization research. We further present design recommendations for future versions of the Reflective Spring Cleaning concept that can be leveraged in other PI systems where software personalization may help refine user behavior.

6.1 Targeted infrequent reflection could better engage less intrinsically motivated users

Prior research has shown that participants who are not intrinsically motivated to self-track, such as those that need to manage medical conditions like diabetes, often quickly become discouraged and stop using PI tools [42, 73, 90]. The relatively recent lived-informatics model of PI has already started to frame lapses in tracking as a normal and positive stage of self-tracking workflow if it aligns with participant motivations [20, 23, 25, 42, 73, 83]. For example, participants may happily disengage with a self-tracking tool after a few weeks of tracking that satisfies their curiosity and helped them explore the problem [21]. However, by contrast it could be a negative experience if the user does not find the reflection helped or was too cumbersome to complete and led to a feeling of failure [21, 73]. We saw that despite being relatively engaged to manage notifications, our participants only felt intrinsic motivation to reflect very infrequently. For example, many participants described that a very infrequent reflection session during Spring Cleaning provided utilitarian value to understand accumulated changes in their notification use. Our findings suggest that future PI research, beyond notification management, could even frame lapses in tracking between targeted reflection sessions as a design goal that may align the intensity of self-tracking with the motivations of more general users who do not have strong intrinsic motivation to frequently self-track [21, 42, 73]. For example, future web browsers might explore adding a personal informatics element that prompts users to engage in a yearly five to ten minute reflection session targeted to help users quickly understand what personal data is being gathered about them, or who has access. Addressing the privacy concerns around extensive self-tracking is an on-going challenge in personal informatics literature, which might help motivate such a system [20].

6.2 Associating PI with software personalizations enhances design goals for both PI and personalization

Prior research has identified two major issues in PI and personalization literature: 1. Users of personalization systems often lack reasons or motivation to overcome barriers to refining their settings [8, 78, 87]. 2. Users who approach reflection from the perspective of behavior change often want clear actionable takeaways from their reflection in order to refine their behavior, which meta studies have suggested are often hard to identify [20, 83]. Our findings show that the novel combination of PI and personalization effectively makes complementary progress on these issues. Reflection on notifications using our data abstractions revealed the surprising and unconscious

costs of distraction to our users, which provided a trigger to personalize [8]. Meanwhile, the ability to personalize notification settings to address the issues identified through reflection provided clear takeaways to reflection. Future software personalization systems should explore ways to visualize whatever behavioral data they already collect in order to help guide personalization decisions. For example, recent online learning platforms have begun to leverage student data to allow educators to analyze and personalize learning activities to suit their students' preferred learning styles [2, 5]. We see opportunities to open this data up to students themselves so they can track and analyze how they best learn new material and perhaps see recommendations for additional activities tailored to those preferences. Further, we showed that 5-10 minutes of reflection on simplified PI did not appear to be a substantial barrier to conducting a personalization session. Though we did not assess precisely how frequently reflection should be done, our data suggests that reflection would be done relatively infrequently, such as at the start of term to identify what courses students might find the most challenging.

6.3 Notification Spring cleaning could be incorporated as a part of holistic wellness systems

Many participants described engaging with Spring Cleaning to reduce unnecessary distraction and improve productivity. Recent research has identified that tools should not focus on a narrow conception of productivity but rather embrace the more holistic realities of how modern work must interweave work and personal tasks for wellness [4, 32, 54, 89]. The stress caused by distraction [47] was a key motivating factor behind our work and several of our participants wanted to manage notifications to reduce stress. They described appreciating how using Spring Cleaning pushed them to establish clearer boundaries against work messages received in personal time. The potential for targeted notification personalization to encourage such reflection suggests that personal informatics systems that support users to track and refine their mental health might benefit from a notification Spring Cleaning feature [50, 57, 59]. Future work could explore how users of these tools, who might be more concerned about maintaining wellness than general notification management, might differ in how they personalize with Spring Cleaning. For example, a few of our participants described how receiving a lot of personal chat messages at work was helpful and positive as long as it occurred during appropriate breaks, such as during lunch. Expanding the participant pool to target users who already track and reflect on their holistic well-being could clarify the extent wellness considerations influence data interpretation. Further, if a short Spring Cleaning session provides interesting and useful insights for users who already self-track, those tools might suggest integrating notification data into their regular wellness tracking habits. Prior work in health and wellness systems have used various forms of guided self-experimentation to help users explore what data to integrate into their self-tracking systems, though this can be a time intensive and difficult process [44, 45, 49]. Further work is needed to explore the extent to which a Spring Cleaning session as a prelude to more intensive tracking might compare to these self-experimentation systems [20].

6.4 Opportunities remain to refine Reflective Spring Cleaning

Prior work has explored a variety of approaches to mixed-initiative notification management that all come with many design trade-offs when compared to Spring Cleaning [6, 14, 34, 53]. For example, notification management suggestions like Prefminer cause distraction from the relatively frequent suggestions themselves and do not contextualize the suggestions with notification data [13, 53]. Suggestion rationales have been explored in other domains to help provide this context, but aren't read by many users [14]. Comparison was supported by the study design, but our data suggests the Reflective Spring Cleaning design may be an effective alternative to traditional mixed-initiative suggestions for notification; participants personalized an average of 28% of their notifications and the majority of participants reported improved experience. However, there is room to refine the Reflective Spring Cleaning design and future comparative work should continue to explore how reflection best fits into notification management workflows.

About half of our participants independently created their own personalizations, often directly inspired by insights gained while reflecting. However, study participation involved some training on how to personalize notifications, and everyday users may be less able to make such personalizations themselves. Design revisions to Spring Cleaning should enable relatively easy independent personalization, such as directly linking the visualizations to manual notification personalization. The easiest implementation could be a simple clickable link anywhere apps are explicitly referenced in the visualizations to the appropriate settings panel for those apps. For example, if users click the bar for Twitter in the "What" facet, the system will open up the Twitter notification settings panel. Linking the visualizations to current phone settings would reduce the knowledge barriers to finding the appropriate settings panels, even if the more complex personalizations such as the rules-based do-not-disturb feature used to implement our suggestions would need additional support. Even more refined versions of Reflective Spring Cleaning would enable users to directly personalize notifications from the graphs. For example, Reflective Spring Cleaning could offer the ability to select bars in the graphs to define a set of notifications to personalize, then assign a rule for how they should be delivered such as highlight, silence or batched delivery (without having to do this in a settings panel).

7 LIMITATIONS

The main limitation of our work is that the prototype was relatively low fidelity. While our prototyping approach allowed us to focus on the design of the reflective components, the suggestions were not overly personalized to each user. The second limitation in our method is that participants may have generated additional insights by self-reflecting in order to answer our interview questions. We were careful to observe participants' initial reactions to their data without further prompting. However, discussing what they learned during reflection in the interview may have impacted participants' personalization decisions. We acknowledge that participants in the wild would not be explicitly prompted to engage in this higher-level reflection of their takeaways from the visualizations. We've noted that our convenience sampling of participants was intentionally

limited to those who are relatively engaged in notification management. Future work remains to extend our sample to a wider population to explore how different backgrounds and motivations might impact how users react to Reflective Spring Cleaning. Adding member checking to the study protocol will also enhance validity.

8 CONCLUSION

In this paper we explored a novel approach to notification management called Reflective Spring Cleaning. Rather than prior mixed-initiative systems that interrupt users with suggestions for notification management, Reflective Spring Cleaning infrequently engages users in a longer reflection and personalization session contextualized by visualizations of notification usage data to encourage reflection on the impact of the notifications they receive. We explored Reflective Spring Cleaning in a four week qualitative study where participants passively tracked their notifications, personalized with the Spring Cleaning prototype, and later reflected on how their notification experience changed. Participants often critically reflected on their notification usage which encouraged many to personalize and address issues identified in their notification management practices. Spring Cleaning improved the notification experience of the majority, often by reducing stress or distraction. Our design concept includes a novel “who-what-when” data abstraction that simplifies and contextualizes key variation within a user’s notifications. We found this data abstraction helped to guide participants’ reflection and supported different user motivations to reflect. We discuss various design implications for future personal informatics and personalization tools. For example, the combination of personalization with personal informatics helped give users clear actionable takeaways from reflection, which has been a challenge in other personal informatics tools. We suggest that similar approaches could be explored in other personal informatics systems, such as a wellness tool prompting users to very infrequently reflect on how they use their mobile devices. Reflective Spring Cleaning takes a novel approach to notification management which continues to be a pressing issue in both academic and industry discourse. Our work brings us an important step forward in empowering the millions of people impacted by disruptive notifications to manage where their attention is directed.

REFERENCES

- [1] Naseem Ahmadpour and Karen Anne Cochrane. 2017. From information to reflection-design strategies for personal informatics. In *2017 IEEE Life Sciences Conference (LSC)*. IEEE, New York, NY, USA, 103–106. <https://doi.org/10.1109/LSC.2017.8268154>
- [2] Hamdan A Alamri, Sunnie Watson, and William Watson. 2021. Learning Technology Models that Support Personalization within Blended Learning Environments in Higher Education. *TechTrends* 65, 1 (2021), 62–78.
- [3] Alyssa E Andrews, Raj M Ratwani, and J Gregory Trafton. 2009. The effect of alert type to an interruption on primary task resumption. In *Proceedings of the HFES Annual Meeting 2009*. HFES, Washington, DC, USA, 1–5.
- [4] Julie Aranda, Noor Ali-Hasan, and Safia Baig. 2016. I’m Just Trying to Survive: An Ethnographic Look at Mobile Notifications and Attention Management. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Florence, Italy) (*Mobile-HCI '16*). Association for Computing Machinery, New York, NY, USA, 564–574. <https://doi.org/10.1145/2957265.2957274>
- [5] Sara Assami, Najima Daoudi, and Rachida Ajhoun. 2018. Personalization criteria for enhancing learner engagement in MOOC platforms. In *2018 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, New York, NY, USA, 1265–1272. <https://doi.org/10.1109/EDUCON.2018.8363375>
- [6] Jonas Auda, Dominik Weber, Alexandra Voit, and Stefan Schneegass. 2018. Understanding User Preferences towards Rule-Based Notification Deferral. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI EA '18*). Association for Computing Machinery, New York, NY, USA, Article LBW584, 6 pages. <https://doi.org/10.1145/3170427.3188688>
- [7] Amid Ayobi, Paul Marshall, and Anna L. Cox. 2016. Reflections on 5 Years of Personal Informatics: Rising Concerns and Emerging Directions. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI EA '16*). Association for Computing Machinery, New York, NY, USA, 2774–2781. <https://doi.org/10.1145/2851581.2892406>
- [8] Nikola Banovic, Fanny Chevalier, Tovi Grossman, and George Fitzmaurice. 2012. Triggering Triggers and Burying Barriers to Customizing Software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (*CHI '12*). Association for Computing Machinery, New York, NY, USA, 2717–2726. <https://doi.org/10.1145/2207676.2208666>
- [9] Eric P.S. Baumer. 2015. Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 585–594. <https://doi.org/10.1145/2702123.2702234>
- [10] Eric P.S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) (*DIS '14*). Association for Computing Machinery, New York, NY, USA, 93–102. <https://doi.org/10.1145/2598510.2598598>
- [11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [12] Andrea Bunt, Cristina Conati, and Joanna McGrenere. 2007. Supporting Interface Customization Using a Mixed-Initiative Approach. In *Proceedings of the 12th International Conference on Intelligent User Interfaces* (Honolulu, Hawaii, USA) (*IUI '07*). Association for Computing Machinery, New York, NY, USA, 92–101. <https://doi.org/10.1145/1216295.1216317>
- [13] Andrea Bunt, Cristina Conati, and Joanna McGrenere. 2010. Mixed-initiative interface personalization as a case study in usable AI. *AI Magazine* 30, 4 (2010), 58.
- [14] Andrea Bunt, Joanna McGrenere, and Cristina Conati. 2007. Understanding the Utility of Rationale in a Mixed-Initiative System for GUI Customization. In *User Modeling 2007*, Cristina Conati, Kathleen McCoy, and Georgios Paliouras (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 147–156.
- [15] Luigi De Russis and Alberto Monge Roffarello. 2017. On the Benefit of Adding User Preferences to Notification Delivery. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI EA '17*). Association for Computing Machinery, New York, NY, USA, 1561–1568. <https://doi.org/10.1145/3027063.3053160>
- [16] Fiona Draxler, Christina Schneegass, Jonas Safranek, and Heinrich Hussmann. 2021. Why Did You Stop? - Investigating Origins and Effects of Interruptions during Mobile Language Learning. In *Mensch Und Computer 2021*. Association for Computing Machinery, New York, NY, USA, 21–33. <https://doi.org/10.1145/3473856.3473881>
- [17] Laura L Ellingson. 2009. *Engaging crystallization in qualitative research: An introduction*. Sage, Newbury Park, CA, USA.
- [18] Chris Elsdén and David S. Kirk. 2014. A Quantified Past: Remembering with Personal Informatics. In *Proceedings of the 2014 Companion Publication on Designing Interactive Systems*. Association for Computing Machinery, New York, NY, USA, 45–48. <https://doi.org/10.1145/2598784.2602778>
- [19] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming Data Complexity in Lifelogs: Exploring Visual Cuts of Personal Informatics Data. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) (*DIS '14*). Association for Computing Machinery, New York, NY, USA, 667–676. <https://doi.org/10.1145/2598510.2598558>
- [20] Daniel A. Epstein, Clara Caldeira, Mayara Costa Figueiredo, Xi Lu, Lucas M. Silva, Lucretia Williams, Jong Ho Lee, Qingyang Li, Simran Ahuja, Qiuer Chen, Payam Dowlatyari, Craig Hilby, Sazeda Sultana, Elizabeth V. Eikey, and Yunan Chen. 2020. Mapping and Taking Stock of the Personal Informatics Literature. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 126 (dec 2020), 38 pages. <https://doi.org/10.1145/3432231>
- [21] Daniel A. Epstein, Monica Caraway, Chuck Johnston, An Ping, James Fogarty, and Sean A. Munson. 2016. Beyond Abandonment to Next Steps: Understanding and Designing for Life after Personal Informatics Tool Use. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 1109–1113. <https://doi.org/10.1145/2858036.2858045>
- [22] Daniel A. Epstein, Bradley H. Jacobson, Elizabeth Bales, David W. McDonald, and Sean A. Munson. 2015. From “Nobody Cares” to “Way to Go!”: A Design Framework for Social Sharing in Personal Informatics. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*

- (Vancouver, BC, Canada) (CSCW '15). Association for Computing Machinery, New York, NY, USA, 1622–1636. <https://doi.org/10.1145/2675133.2675135>
- [23] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) (*UbiComp '15*). Association for Computing Machinery, New York, NY, USA, 731–742. <https://doi.org/10.1145/2750858.2804250>
- [24] Joel E. Fischer, Chris Greenhalgh, and Steve Benford. 2011. Investigating Episodes of Mobile Phone Activity as Indicators of Opportune Moments to Deliver Notifications. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden) (*MobileHCI '11*). Association for Computing Machinery, New York, NY, USA, 181–190. <https://doi.org/10.1145/2037373.2037402>
- [25] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on Reflection: Framing a Design Landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (Brisbane, Australia) (*OZCHI '10*). Association for Computing Machinery, New York, NY, USA, 216–223. <https://doi.org/10.1145/1952222.1952269>
- [26] fruitranger. 2017. Notification Manager - Cleaner. <https://play.google.com/store/apps/details?id=com.cortana.notificationmanager.notificationclean&hl=en>. Accessed: 2020-3-11.
- [27] Krzysztof Z. Gajos, Mary Czerwinski, Desney S. Tan, and Daniel S. Weld. 2006. Exploring the Design Space for Adaptive Graphical User Interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (Venezia, Italy) (*AVI '06*). Association for Computing Machinery, New York, NY, USA, 201–208. <https://doi.org/10.1145/1133265.1133306>
- [28] Krzysztof Z. Gajos, Katherine Everitt, Desney S. Tan, Mary Czerwinski, and Daniel S. Weld. 2008. Predictability and Accuracy in Adaptive User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) (*CHI '08*). Association for Computing Machinery, New York, NY, USA, 1271–1274. <https://doi.org/10.1145/1357054.1357252>
- [29] Larry Goldberg, Bettijoyce Lide, Svetlana Lowry, Holly A Massett, Trisha O'Connell, Jennifer Preece, Whitney Quesenbery, and Ben Shneiderman. 2011. Usability and accessibility in consumer health informatics: current trends and future challenges. *American journal of preventive medicine* 40, 5 (2011), S187–S197.
- [30] Google. 2021. Google Play Store. <https://play.google.com/store/apps>. Accessed: 2021-09-1.
- [31] Tom Gross and Anna-Lena Mueller. 2021. NotificationManager: Personal Boundary Management on Mobile Devices. In *Human-Computer Interaction - INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 243–261.
- [32] Hayley Guillou, Kevin Chow, Thomas Fritz, and Joanna McGrenere. 2020. Is Your Time Well Spent? Reflecting on Knowledge Work More Holistically. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3313831.3376586>
- [33] Rebecca Gulotta, Jodi Forlizzi, Rayoung Yang, and Mark Wah Newman. 2016. Fostering Engagement with Personal Informatics Systems. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) (*DIS '16*). Association for Computing Machinery, New York, NY, USA, 286–300. <https://doi.org/10.1145/2901790.2901803>
- [34] Brian Heater and Sarah Perez. 2021. Apple refines iOS 15 notifications with Focus, Summary features. <https://techcrunch.com/2021/06/07/apple-refines-ios-15-notifications-with-focus-summary-features/>. Accessed: 2021-09-01.
- [35] Jerry Hildenbrad. 2017. Android Oreo will make you love notifications again. <https://www.androidcentral.com/android-o-and-notifications-everything-you-need-know>. Accessed: 2017-11-10.
- [36] Helen M. Hodgetts and Dylan M. Jones. 2007. Reminders, Alerts and Pop-ups: The Cost of Computer-Initiated Interruptions. In *Human-Computer Interaction. Interaction Design and Usability*, Julie A. Jacko (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 818–826.
- [37] Eric Horvitz, Johnson Apacible, and Muru Subramani. 2005. Balancing awareness and interruption: Investigation of notification deferral policies. *User Modeling 2005* 1, 1 (2005), 149–149.
- [38] Roberto Hoyle, Srijita Das, Apu Kapadia, Adam J. Lee, and Kami Vaniea. 2017. Was My Message Read? Privacy and Signaling on Facebook Messenger. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 3838–3842. <https://doi.org/10.1145/3025453.3025925>
- [39] Elle Hunt. 2020. One ping after another: why everyone needs a notification detox. <https://www.theatlantic.com/technology/archive/2015/07/push-notifications-versus-phone-calls/398081/>. Accessed: 2021-09-01.
- [40] 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* (Savannah, Georgia, USA) (*CSCW '10*). Association for Computing Machinery, New York, NY, USA, 27–30. <https://doi.org/10.1145/1718918.1718926>
- [41] Izabelle Janzen, Francesco Vitale, and Joanna McGrenere. 2018. Control and Personalization: Younger versus Older Users' Experience of Notifications. In *Proceedings of Graphics Interface 2018* (Toronto, Ontario) (*GI 2018*). Canadian Human-Computer Communications Society / Société canadienne du dialogue humain-machine, Toronto, ON, Canada, 129 – 136. <https://doi.org/10.20380/GI2018.17>
- [42] Mohammad Hossein Jarrahi, Nicci Gafinowitz, and Grace Shin. 2018. Activity trackers, prior motivation, and perceived informational and motivational affordances. *Personal and Ubiquitous Computing* 22, 2 (2018), 433–448.
- [43] Simon L. Jones and Ryan Kelly. 2018. Dealing With Information Overload in Multifaceted Personal Informatics Systems. *Human-Computer Interaction* 33, 1 (2018), 1–48. <https://doi.org/10.1080/07370024.2017.1302334>
- [44] Ravi Karkar, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, Roger Vilardaga, and Jasmine Zia. 2017. TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 6850–6863. <https://doi.org/10.1145/3025453.3025480>
- [45] Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A. Kientz. 2012. Lullaby: A Capture & Access System for Understanding the Sleep Environment. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (Pittsburgh, Pennsylvania) (*UbiComp '12*). Association for Computing Machinery, New York, NY, USA, 226–234. <https://doi.org/10.1145/2370216.2370253>
- [46] Vera Khovanskaya, Eric P.S. Baumer, Dan Cosley, Stephen Volda, and Geri Gay. 2013. "Everybody Knows What You're Doing": A Critical Design Approach to Personal Informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (*CHI '13*). Association for Computing Machinery, New York, NY, USA, 3403–3412. <https://doi.org/10.1145/2470654.2466467>
- [47] Kostadin Kushlev and Elizabeth W Dunn. 2015. Checking email less frequently reduces stress. *Computers in Human Behavior* 43 (2015), 220–228.
- [48] Pat Langley. 1997. Machine learning for adaptive user interfaces. In *KI-97: Advances in artificial intelligence*. Springer, Springer, Berlin, Heidelberg, 53–62.
- [49] Jisoo Lee, Erin Walker, Winslow Burleson, Matthew Kay, Matthew Buman, and Eric B. Heckler. 2017. Self-Experimentation for Behavior Change: Design and Formative Evaluation of Two Approaches. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 6837–6849. <https://doi.org/10.1145/3025453.3026038>
- [50] Kwangyoung Lee and Hwajung Hong. 2018. MindNavigator: Exploring the Stress and Self-Interventions for Mental Wellness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3173574.3174146>
- [51] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding My Data, Myself: Supporting Self-Reflection with Ubicomp Technologies. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (Beijing, China) (*UbiComp '11*). Association for Computing Machinery, New York, NY, USA, 405–414. <https://doi.org/10.1145/2030112.2030166>
- [52] Deborah Lupton. 2014. Self-Tracking Cultures: Towards a Sociology of Personal Informatics. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design* (Sydney, New South Wales, Australia) (*OzCHI '14*). Association for Computing Machinery, New York, NY, USA, 77–86. <https://doi.org/10.1145/2686612.2686623>
- [53] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. 2016. PrefMiner: Mining User's Preferences for Intelligent Mobile Notification Management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) (*UbiComp '16*). Association for Computing Machinery, New York, NY, USA, 1223–1234. <https://doi.org/10.1145/2971648.2971747>
- [54] 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*. Association for Computing Machinery, New York, NY, USA, 1021–1032. <https://doi.org/10.1145/2858036.2858566>
- [55] Robinson Meyer. 2015. Push Notifications Are as Distracting as Phone Calls. <https://www.theguardian.com/lifeandstyle/2020/jan/27/one-ping-after-another-how-constant-notifications-are-driving-us-to-distraction>. Accessed: 2021-09-01.
- [56] Jack Mezirow. 1998. On critical reflection. *Adult education quarterly* 48, 3 (1998), 185–198.
- [57] Sean A Munson. 2017. Rethinking assumptions in the design of health and wellness tracking tools. *Interactions* 25, 1 (2017), 62–65.
- [58] Tamara Munzner. 2009. A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics* 15, 6 (2009), 921–928.
- [59] Elizabeth L. Murnane, Tara G. Walker, Beck Tench, Stephen Volda, and Jaime Snyder. 2018. Personal Informatics in Interpersonal Contexts: Towards the Design of Technology That Supports the Social Ecologies of Long-Term Mental Health

- Management. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 127 (nov 2018), 27 pages. <https://doi.org/10.1145/3274396>
- [60] Midas Nouwens, Carla F. Griggio, and Wendy E. Mackay. 2017. "WhatsApp is for Family; Messenger is for Friends": Communication Places in App Ecosystems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 727–735. <https://doi.org/10.1145/3025453.3025484>
- [61] Anna Offenwanger, Alan John Milligan, Minsuk Chang, Julia Bullard, and Dongwook Yoon. 2021. Diagnosing Bias in the Gender Representation of HCI Research Participants: How It Happens and Where We Are. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (*CHI '21*). Association for Computing Machinery, New York, NY, USA, Article 399, 18 pages. <https://doi.org/10.1145/3411764.3445383>
- [62] Fredrik Ohlin and Carl Magnus Olsson. 2015. Intelligent Computing in Personal Informatics: Key Design Considerations. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (Atlanta, Georgia, USA) (*IUI '15*). Association for Computing Machinery, New York, NY, USA, 263–274. <https://doi.org/10.1145/2678025.2701378>
- [63] Fredrik Ohlin, Carl Magnus Olsson, and Paul Davidsson. 2015. Analyzing the Design Space of Personal Informatics: A State-of-practice Based Classification of Existing Tools. In *Universal Access in Human-Computer Interaction. Access to Today's Technologies*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 85–97.
- [64] Tadashi Okoshi, Jin Nakazawa, and Hideyuki Tokuda. 2016. Interruptibility Research: Opportunities for Future Flourishment. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (Heidelberg, Germany) (*UbiComp '16*). Association for Computing Machinery, New York, NY, USA, 1524–1529. <https://doi.org/10.1145/2968219.2968543>
- [65] 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* (Seattle, Washington) (*UbiComp '14*). Association for Computing Machinery, New York, NY, USA, 897–908. <https://doi.org/10.1145/2632048.2632062>
- [66] 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* (Toronto, ON, Canada) (*MobileHCI '14*). Association for Computing Machinery, New York, NY, USA, 233–242. <https://doi.org/10.1145/2628363.2628364>
- [67] 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* (Toronto, Ontario, Canada) (*CHI '14*). Association for Computing Machinery, New York, NY, USA, 3319–3328. <https://doi.org/10.1145/2556288.2556973>
- [68] Afarin Pirzadeh, Li He, and Erik Stolterman. 2013. Personal Informatics and Reflection: A Critical Examination of the Nature of Reflection. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1979–1988. <https://doi.org/10.1145/2468356.2468715>
- [69] Romina Poguntke, Christina Schneegass, Lucas Van der Vekens, Rufat Rzyayev, Jonas Auda, Stefan Schneegass, and Albrecht Schmidt. 2020. NotiModes: An Investigation of Notification Delay Modes and Their Effects on Smartphone Users. In *Proceedings of the Conference on Mensch Und Computer* (Magdeburg, Germany) (*MuC '20*). Association for Computing Machinery, New York, NY, USA, 415–419. <https://doi.org/10.1145/3404983.3410006>
- [70] James O Prochaska and Wayne F Velicer. 1997. The transtheoretical model of health behavior change. *American journal of health promotion* 12, 1 (1997), 38–48.
- [71] Aare Puussaar, Adrian K. Clear, and Peter Wright. 2017. Enhancing Personal Informatics Through Social Sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 6936–6942. <https://doi.org/10.1145/3025453.3025804>
- [72] Amon Rapp and Federica Cena. 2014. Self-monitoring and Technology: Challenges and Open Issues in Personal Informatics. In *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*, Constantine Stephanidis and Margherita Antona (Eds.). Springer International Publishing, Cham, 613–622.
- [73] Amon Rapp and Federica Cena. 2016. Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human-Computer Studies* 94 (2016), 1–17. <https://doi.org/10.1016/j.ijhcs.2016.05.006>
- [74] Amon Rapp, Alessandro Marcengo, Luca Buriano, Giancarlo Ruffo, Mirko Lai, and Federica Cena. 2018. Designing a personal informatics system for users without experience in self-tracking: a case study. *Behaviour & Information Technology* 37, 4 (2018), 335–366. <https://doi.org/10.1080/0144929X.2018.1436592>
- [75] Amon Rapp and Maurizio Tirassa. 2017. Know Thyself: A Theory of the Self for Personal Informatics. *Human-Computer Interaction* 32, 5-6 (2017), 335–380. <https://doi.org/10.1080/07370024.2017.1285704>
- [76] Raj M. Ratwani, Alyssa E. Andrews, Jenny D. Sousek, and J. Gregory Trafton. 2008. The Effect of Interruption Modality on Primary Task Resumption. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 52, 4, 393–397. <https://doi.org/10.1177/154193120805200441>
- [77] RescueTime. 2021. RescueTime. <https://www.rescuetime.com/>. Accessed: 2021-09-1.
- [78] 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* (Toronto, Ontario, Canada) (*CHI '14*). Association for Computing Machinery, New York, NY, USA, 3055–3064. <https://doi.org/10.1145/2556288.2557189>
- [79] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. 2005. Reflective Design. In *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility* (Aarhus, Denmark) (*CC '05*). Association for Computing Machinery, New York, NY, USA, 49–58. <https://doi.org/10.1145/1094562.1094569>
- [80] MobileSoft s.r.o. 2020. AppBlock - Stay Focused (Block Websites and Apps). <https://play.google.com/store/apps/details?id=cz.mobilesoft.appblock&hl=en>. Accessed: 2020-3-11.
- [81] Timothy J Trull and Ulrich W Ebner-Priemer. 2009. Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. *Psychological assessment* 21, 4 (2009), 5 pages.
- [82] Liam D. Turner, Stuart M. Allen, and Roger M. Whitaker. 2015. Interruptibility Prediction for Ubiquitous Systems: Conventions and New Directions from a Growing Field. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) (*UbiComp '15*). Association for Computing Machinery, New York, NY, USA, 801–812. <https://doi.org/10.1145/2750858.2807514>
- [83] Elisabeth T. Kersten van Dijk, Joyce H.D.M. Westerink, Femke Beute, and Wijnand A. IJsselstein. 2017. Personal Informatics, Self-Insight, and Behavior Change: A Critical Review of Current Literature. *Human-Computer Interaction* 32, 5-6 (2017), 268–296. <https://doi.org/10.1080/07370024.2016.1276456>
- [84] Alexandra Voit, Dominik Weber, and Niels Henze. 2018. Qualitative Investigation of Multi-Device Notifications. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) (*UbiComp '18*). Association for Computing Machinery, New York, NY, USA, 1263–1270. <https://doi.org/10.1145/3267305.3274117>
- [85] 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* (Barcelona, Spain) (*MobileHCI '18*). Association for Computing Machinery, New York, NY, USA, Article 2, 13 pages. <https://doi.org/10.1145/3229434.3229436>
- [86] Dominik Weber, Alexandra Voit, and Niels Henze. 2018. Notification Log: An Open-Source Framework for Notification Research on Mobile Devices. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) (*UbiComp '18*). Association for Computing Machinery, New York, NY, USA, 1271–1278. <https://doi.org/10.1145/3267305.3274118>
- [87] 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* (Hamburg, Germany) (*MuC'19*). Association for Computing Machinery, New York, NY, USA, 361–372. <https://doi.org/10.1145/3340764.3340765>
- [88] Dominik Weber, Alexandra Voit, Huy Viet Le, and Niels Henze. 2016. Notification Dashboard: Enabling Reflection on Mobile Notifications. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Florence, Italy) (*MobileHCI '16*). ACM, New York, NY, USA, 936–941. <https://doi.org/10.1145/2957265.2962660>
- [89] Joan C Williams, Jennifer L Berdahl, and Joseph A Vandello. 2016. Beyond work-life "integration". *Annual review of psychology* 67 (2016), 515–539.
- [90] Nan Yang, Gerbrand van Hout, Loe Feijs, Wei Chen, and Jun Hu. 2020. Enhancing intrinsic motivation in physical activity through quantified-self data sharing. *EAI Endorsed Transactions on Pervasive Health and Technology* 6, 21 (2020), e4.

A TARGETED SURVEY RESULTS

Table 2 contains the list of 20 suggestions given to participants in the initial targeted survey. From these 20, the 7 suggestions had their wording simplified to be used in the reflection study and are shown in table 1.

Table 2: The 20 Notification Management Suggestions given in the targeted survey with rates of acceptance.

Acceptance Rate	Suggestion Text
82.75	Automatically turn notification delivery to silent during scheduled meetings on your calendar.
82.14	Highlight work emails that need you to respond to them within a day with a special ringtone or vibration.
79.31	Set working hours (e.g. 9am to 5pm), where less important non-work notifications (e.g. social media activity or messages from non-close friends) are silent.
78.57	Ask once if you want to unsubscribe from any email advertisements.
75.00	Highlight messages from important contacts (e.g. your parents partner, spouse, children or very close friends) with a special ringtone or vibration.
71.42	Deliver system messages silently
67.85	Deliver social media activity notifications (e.g. status updates, not chat messages) that are not directly related to, or reference you silently.
64.28	Highlight all work or school related email with a special ringtone or vibration.
64.28	Silently deliver email advertisements.
57.14	Set off-work hours (e.g. after 6pm, Monday to Friday) where work notifications are silent.
57.14	Deliver social media activity notifications (e.g. status updates but not chat messages) silently at all times.
53.57	Provide a digest notification of important meetings from your calendar that day in the morning.
39.28	Deliver broadcast style work messages silently.
39.28	Highlight direct messages from chat applications (not including group chats) with a special sound or vibration.
35.71	Deliver social media activity notifications (e.g. status updates, but not chat messages) in a batch every hour.
32.14	Deliver notifications for group chat messages in a batch every hour.
28.57	Deliver social media notifications (including chat messages) in a batch every hour.
28.57	Do not deliver social media notifications (including chat messages) during the day. Instead deliver a batch of everything you received during the day in the evening (e.g. at 6pm).
28.57	Do not deliver notifications from video games during the day. Instead deliver a batch of the ones you received during the day in the evening.
28.57	Deliver all notifications from video games silently at all times.

B CATEGORIES OF APP FROM WHAT ABSTRACTION

Table 3: The 13 categories of app presented in the what abstraction.

Category	Rough Description
System	Software updates and os messages
Social Media	various forms of social media (e.g. twitter, linkedin)
Games	video games and other forms of entertainment
Email	Any email notifications
Text	Any direct messaging (e.g. discord, slack)
Voice	Direct phone calls and video conferencing
Calendar	Reminders and calendar events
News	News reading apps
Tools	Productivity tools (e.g. google drive or slides)
Market	Purchasing, sales or apps for specific stores
Other	Catchall other