Annotif: A System for Annotating Mobile Notifications in User Studies

Dominik Weber
University of Stuttgart
Stuttgart, Germany
weberdk@vis.uni-stuttgart.de

Alexandra Voit
University of Stuttgart
Stuttgart, Germany
info@alexandra-voit.de

Gisela Kollotzek
University of Stuttgart
Stuttgart, Germany
gk-stuttgart@web.de

Niels Henze
University of Regensburg
Regensburg, Germany
niels.henze@ur.de

ABSTRACT

Notifications are an essential feature of smartphones. While they support users in staying up-to-date, they are also a prominent source of interruptions. A deeper understanding of mobile notifications is required to avoid adverse effects. However, assessing mobile notifications is challenging as user studies on mobile notifications are typically conducted in-situ. Surveying users may lead to additional interruptions, and the content of notifications is inherently private. In this paper, we introduce a privacy-aware system for annotating mobile notifications in user studies. In an in-situ case study, participants annotated their notifications for one week. Participants perceived 38.91% of their notifications as not important and over half (51.75%) as non-urgent. Only 6.33% of the notifications were rated as both very important and very urgent. We discuss influencing factors, including a breakdown of messaging notifications, and implications for future smart notification systems that continue to fulfill users' information need while respecting their digital well-being.

CCS CONCEPTS

• **Human-centered computing → User studies**; **Field studies**; **Empirical studies in HCI**.

KEYWORDS

Mobile notifications, annotation, in-situ, in-the-wild, smartphones, interruptions, importance, urgency.

ACM Reference Format:
Dominik Weber, Alexandra Voit, Gisela Kollotzek, and Niels Henze. 2019. Annotif: A System for Annotating Mobile Notifications in User Studies. In *MUM 2019: 18th International Conference on Mobile and Ubiquitous Multimedia (MUM 2019), November 26–29, 2019, Pisa, Italy*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3365610.3365611

MUM 2019, November 26–29, 2019, Pisa, Italy

ACM ISBN 978-1-4503-7624-2/19/11…$15.00
https://doi.org/10.1145/3365610.3365611

1 INTRODUCTION

Notifications are an integral part of current smartphones. Apps can trigger notifications to gain the user's attention using visual, tactile, and auditory cues. These notifications can be of different categories, from communication, events, news, utilities to system alerts. Prior work found that users value notifications related to people and events [37]. Instant messaging notifications enable asynchronous communication and allow people to stay connected [8]. However, while these types of notifications are valued, they can also cause interruptions and distractions. This can cause adverse effects, such as increased stress [48], inattention [25], and reduced task performance [10]. Prior work suggests to reduce adverse effects by silencing [8], filtering [27], or deferring [30] notifications. This is a balancing act of fulfilling the users' information need while respecting their digital well-being. A "smart" notification management system should not withhold important and time-sensitive notifications from users.

A deeper understanding of mobile notifications is still missing. For instance, while previous work found that "notifications are for messaging" [37], communication is a broad category which requires closer inspection. In a recently published study, researchers found that participants attend notifications about individual (1:1) chats faster than group chats [36]. Further, messaging is no longer limited to text messages alone. Modern messaging apps support rich media formats such as pictures, videos, and voice recordings. In July 2017, the popular messaging app *WhatsApp* reported 1 billion daily active users sending 55 billion messages, 4.5 billion photos, and 1 billion videos per day [47]. It is not yet known how these types of messages, and corresponding notifications, are perceived by users.

Learning how users perceive notifications is a challenging task. The importance and urgency of notifications depend on their content and context [28], which is the reason why notifications are typically assessed in-situ. In the case of communication-related notifications, the relationship between the sender and the user matters as well [29]. The *Experience Sampling Method* (ESM) is a popular method to learn about users' experience throughout the day by triggering questionnaires at random times or using specific events [6, 20]. However, these questionnaires are delivered to users as notifications themselves, which makes it difficult to survey users without influencing them by introducing additional interruptions. Further, the nature of notifications is inherently private and often sensitive. While the content of notifications is essential to understand the perceived importance and urgency, handling the content in user studies must not be an afterthought.

In this paper, we introduce *Annotif*, a system for annotating mobile notifications in user studies. Using this system, we conducted a week-long in-situ case study to explore the importance and urgency of mobile notifications. *Annotif* enabled participants to annotate their notifications including the content and context while respecting the notifications' private nature. The results show that participants perceived 38.91% of their notifications as not important and over half (51.75%) as non-urgent. Only 6.33% of the notifications were rated as both very important and very urgent. We discuss influencing factors, including a detailed breakdown of 1:1 and group messaging notifications.

Our contribution is two-fold: (1) We introduce *Annotif*, a privacy-aware system for annotating mobile notifications in in-situ studies in an unobtrusive manner. (2) We report the results of a week-long in-situ case study in which participants used *Annotif* to annotate the importance and urgency of their notifications.

2 RELATED WORK

In this section, we summarize related work on mobile notifications, interruptions, and experience sampling.

Mobile Notifications

A body of prior work investigated mobile notifications. The smartphone is currently the preferred device to be notified on [45]. Sahami Shirazi et al. conducted a large-scale assessment of mobile notifications with almost 200 million notifications from over 40,000 users [37, 41]. They found that notifications from messaging apps are regarded as highly valuable. In general, the researchers argue that important notifications are about people and events. Pielot et al. investigated mobile notifications in a smaller, more controlled study [33]. They found that their 15 participants received $M = 63.5$ notifications per day in the year 2014. Pielot et al. recently revisited this work in a large-scale study with 278 Spanish users [36]. They found that the participants received $Md = 56$ notifications per day, with most of them being messaging notifications.

Mobile phones enabled text messaging as a popular communication method [5]. Dingler et al. explored the attentiveness of users towards mobile messages [13]. They found that users were attentive to messages for approximately 12 hours per day. Avrahami et al. showed that the responsiveness towards instant messaging is affected by the context and the presentation of messages [4]. Church et al. investigated the effect of the messaging app *WhatsApp* overtaking traditional *SMS* [8]. Mehrotra et al. found that the sender of messages can have an impact on how notifications are perceived [29] and Pielot et al. identified features to predict if a user will attend a message within a specified period [34].

Interruptions and Notification Deferral

Prior studies investigated what makes interruptions disruptive [17]. Interruptions can delay task completion by up to four times [26]. Czerwinski et al. explored the adverse effects of instant messaging interruptions on different kinds of tasks [10]. Other work looked at the attentional cost of receiving notifications [38] and relevant interruptions [11, 16, 18, 23]. While interruptions may cause inattention [25], intense phone use does not predict negative well-being [24]. In a world of constant connection, being unavailable is an interesting research topic [7]. Researchers recently started to explore the "joy of missing out" [2], enabling users to reflect on their notifications [46], and experimented with disabling push notification altogether [35].

Other approaches explored models to better time interruptions [1, 32, 39, 40]. *Attelia II* is a middle-ware to defer notifications to times between subsequent tasks [30]. *SCAN* is another approach of a notification system that takes the social context into account [31]. *PrefMiner* is a system to generate rules for notification management automatically [27]. Fischer et al. investigated mobile phone activity as indicators of opportune moments to deliver notifications [15]. Iqbal and Bailey explored the use of breakpoints to deliver notifications [21]. Weber et al. investigated the user-defined deferral of mobile notifications by "snoozing" them [42], and Auda et al. enabled users to define rules for automatically snoozing notifications and creating end-of-day summaries [3].

Experience Sampling of Mobile Notifications

The importance and urgency of notifications depend on their content and context [28]. The *Experience Sampling Method* (ESM) is a popular method to learn about participants' experience throughout the day by triggering surveys at random times or triggered by specific events [6, 9, 20]. However, these surveys may disrupt participants [19] which is problematic when studying interruptions caused by notifications in the first place. Sahami Shirazi et al. triggered surveys on participants' desktop computer to assess the importance of mobile notifications [37]. However, the surveys were limited to few notifications per participant and only surveyed about the app that created the notification, without considering the content or context. The researchers balanced the limited samples by having a user base of over 40,000 users. However, for most user studies, especially in an academic context, this number of users is unfeasible.

An alternative to ESM is the *Day Reconstruction Method* (DRM) [22]. Instead of surveying participants throughout the day, they are asked to reconstruct the day systematically before assessing it. This method does not capture participants' impressions in the exact moment. Instead, the assessment is done post-hoc. However, by asking participants to reconstruct the day this limitation is reduced, while having the advantage of not disrupting the participants during the day.

Summary

Research on mobile notifications is typically conducted in-situ. Capturing how users perceive notifications without influencing them is a challenging task. Further, while prior work has shown the importance of the notification content, handling the content in user studies is challenging due to its private nature. What is missing are tools that enable us to gain a deeper understanding of mobile notifications in an unobtrusive and privacy-respecting manner.

3 THE "ANNOTIF" ANNOTATION SYSTEM

To overcome the challenge of surveying users about mobile notifications, we developed an annotation system consisting of an Android notification logging app, a server application, and a web-based annotation tool (see Figure 1).

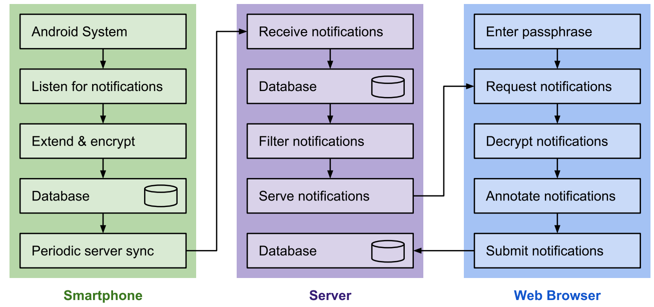

Figure 1: The data flow of the system. Notifications are collected on the smartphone and periodically synced to the server. On the server, notifications are filtered and served to the web browser. The notifications are then annotated by the user and sent back to the server. Notification content is encrypted between the smartphone and the web browser.

Background on Notifications in Android

Notifications are a core feature of Android. Any Android app can trigger notifications by default. However, users can disable notifications for specific apps. A notification typically consists of a small icon and two lines of text. Notifications are shown in the notification drawer that can always be accessed by swiping down from the top of the screen [44]. In newer versions of Android, notifications are also shown on the lock screen. Notifications can be extended in several ways. Developers can attach sounds and vibration patterns, set priority levels, and group multiple notifications.

The *priority* level is one factor that decides the order of notifications in the notification drawer. Possible priority levels range from `MIN`, `LOW`, `DEFAULT`, `HIGH`, to `MAX`. The notifying app sets the priority level.

Notification *groups* consist of multiple notifications from the same app that share the same *group key*. Apps can set one notification as the *group summary*. For example, consider an instant messaging app that creates a notification for each unread conversation. The app would create N notifications for the N unread conversations, and an additional notification as the group summary. Depending on the Android version, the Android system would only display the group summary while hiding the other N notifications, or allow the group summary to be expanded. When it comes to updating existing notifications, apps may use different strategies. In the example with $N + 1$ instant messaging notifications, an additional conversation could cause the app to trigger a single notification for the new conversation and update the group summary, resulting in two notification events. An alternative strategy is to revoke and re-create all notifications, resulting in $N + 2$ notification events. It is important to keep this behavior in mind, as notification events in the Android system do not directly correspond to the actual number of different notifications shown to users.

Notification Logging App

Based on our prior work [43], we developed an Android app to log notifications from smartphones, extend them with context data, encrypt them, and periodically sync them with a server. The app registers itself as a *Notification Listener Service* [12]. The service retrieves events about new and removed notifications from all apps installed on the device. A useful aspect of this service is that it is exempt from battery optimization procedures and therefore always runs in the background without interruption.

The left side of Figure 1 shows the data flow of the Android app. The *Notification Listener Service* listens for new notifications in the background. Once a new notification event is received, the app first extracts all meta-data of the notification. This includes the *package name* (the identifier of the app that created the notification), the time when the notification was created, the app-set priority level, and the notification group key. In line with prior work on smartphone users' concerns [14] and to respect the private nature

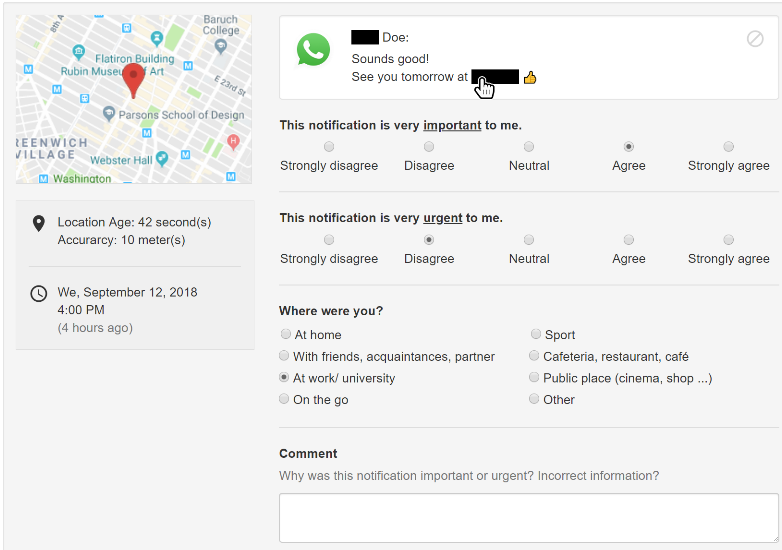

Figure 2: Screenshot of the web-based annotation tool for a single notification. The tool provides a map, the absolute and relative time, the app icon of the notifying app, the notification content text, importance/urgency ratings on a 5-point scale, a location selection, and an optional comment field. Clicking on words censors them. Users can optionally censor the entire notification content by clicking on the block icon in the top right corner.

of notifications, the app extracts the notification content and encrypts it using the *Advanced Encryption Standard* (AES) in the *Cipher Block Chaining* (CBC) mode with an encryption key derived from a user-defined passphrase using `PBKDF2`. The app also stores the `SHA-256` hash of the content. This allows the detection of notifications with duplicate content without knowledge of the content itself. The app will also automatically record if the content contains the name of the user or the name of a user's contact. This is done by retrieving the list of saved contacts on the device and searching for the contact names in the content text using a regular expression. Finally, the app associates the device's current location with the notification. The extended and encrypted notifications are then stored in an on-device database and periodically sent in batches to the server using a secure connection.

Server

The server receives notifications from the Android app (see Figure 1). It stores the encrypted notifications in a database and associates them with the user ID. To prepare the notifications for annotation, the server executes the following five filtering steps:

(1) For a given user ID, select all notifications that were not yet annotated by the user.
(2) Discard all notifications that are older than a specific time period (e.g., 48 hours). This ensures that users only annotate notifications that they still remember.
(3) Cluster the remaining notifications according to the app that created them. Split the clusters if they contain notifications from multiple days to one cluster per day. Sort the clusters from old to new.
(4) For each cluster, retrieve the group key of all notifications. For each group key in each cluster, check if the group contains both a group summary and other notifications. If this is true, filter the group summary and keep the other notifications. Otherwise, if there is only a group summary, keep it.
(5) For all notifications in each cluster, compare the text hash value. If there are multiple notifications from the same app with the same text hash, keep the first instance of the notification and discard the duplicates.

The remaining clusters do not contain notifications from the same app with duplicate content. The clusters are served one-by-one to the web-based annotation tool. For instance, the user would see a list of *WhatsApp* notifications that were created approximately at the same time. After they are annotated, the server stores them in a separate database table. The server would then serve the next cluster of notifications, e.g., a set of email notifications. Users can take a break from annotating notifications at any point in time.

Annotation Tool

Users access the web-based annotation tool using a personalized link that contains the user ID. The user is then shown a password field to enter the same passphrase that was set in the Android app. The passphrase is kept client-side and never sent to the server. The annotation tool then requests a new set of notifications from the server, decrypts them using the user-provided passphrase, and renders them (see Figure 2). The annotation box consists of the following parts:

Location. A map showing the location of the device when the notification was triggered. It also shows the estimated location accuracy and age of the location data.

Date and Time. The date and time when the notification was triggered, including a relative description to the current time (e.g., “4 hours ago”).

Notification. The content of the notification. Clicking on a word censors it. The “block” icon in the top right corner censors all text at once. Additionally, the icon of the app that created the notification is shown next to the text.

Annotation Form Controls. We implemented 5-point Likert scale items to rate the agreement to the statements that the notification is very important/ very urgent. Further, the location can be assigned to one of eight pre-defined labels, and an optional free text field allows users to provide additional information or to report problems.

After the user annotated all notifications in a cluster, the tool verifies that all required form controls were selected and sends the annotated notifications to the server using a secure connection. At this point, the content is no longer encrypted and can be used for analysis. However, users are in control about what is being shared by censoring parts of or the entire content. We want to highlight that the system can be easily modified or extended by logging additional values in the Android app or by replacing the form controls shown in the annotation tool. Notifications are stored in the *JavaScript Object Notation* (JSON) and are extended as they flow through the system. Therefore, the system can be used flexibly in different kinds of user studies.

4 CASE STUDY

To test the *Annotif* system, we conducted a week-long in-situ study. Participants installed the Android logging app on their personal smartphones and annotated notifications on their personal laptops or desktop PCs.

Design

We designed the case study inspired by prior work [37], the *Experience Sampling Method* (ESM), and the *Day Reconstruction Method* (DRM). The Android app on participants' personal smartphones recorded all notifications throughout the day and periodically synced them with the server. The notifications were extended with additional context data, such as the location of the device when the notification was triggered. This allowed the participants to reflect on the context when the notification was received. Participants annotated the notifications on their personal laptops or desktop PCs. They were free to annotate them whenever they had time. However, we set a time limit of 48 hours to ensure that participants could still remember the context.

Procedure

We invited the participants to our lab one-by-one and explained the study procedure. We explicitly stated that the participation is voluntary and that they can end their participation at any time. After the introduction, participants signed a consent form and filled in a demographic survey. We then installed the notification logging app on participants' smartphones. On all devices, we verified that the date and time were set correctly, that there was sufficient free storage available, and that location services were enabled. Participants then entered their first, last, and nicknames in the app. The app also accessed the names of the participants' contacts to automatically detect if a notification contained the name of a contact. After the participants set their secret passphrase for the text encryption, the notification logging app was silently running in the background of the smartphones. No further intervention was necessary.

We then showed the participants the annotation tool and explained all aspects of it. After the participants left, we sent out personalized emails with links to the annotation tool. Participants then annotated notifications for one week. Afterward, we sent out another email with instructions on how to uninstall the app and a post-study questionnaire.

Participants

We recruited participants from the local area. All participants were German. A requirement for the study was that participants had to own an Android-based smartphone with Android 5.0 or newer, and a laptop or desktop PC. Thirteen participants participated in the study (7 female, 6 male). They were between 21 and 55 years old (M=26.23; SD=8.44). Four participants were employees, and nine were students. Ten of the thirteen participants stated to use their smartphones for both personal and work purposes. The other three participants only used their smartphone for personal purposes.

5 RESULTS

All thirteen participants completed the study and annotated their notifications for one week. Participants used their personal smartphones for the study. The devices had between 151 and 391 apps installed (M=266; SD=92). This number includes system apps and apps that were pre-installed by the device manufacturer. The language of all smartphones was set to German.

Table 1: Total notification events, filtered events (duplicates/ groups), missed annotations and annotated notifications.

#	Total events	Filtered	Missed	Annotated
P1	792	539	0	253
P2	1,219	816	31	372
P3	1,630	955	0	675
P4	2,327	1,518	86	723
P5	810	427	99	284
P6	738	458	0	280
P7	1,321	721	0	600
P8	917	611	0	306
P9	594	335	0	259
P10	2,664	1,855	48	761
P11	1,240	745	9	486
P12	1,563	1,099	0	464
P13	2,856	1,940	191	725
$\sum$	18,671	12,019	464	6,188

Received and Annotated Notifications

We logged a total of 18,671 notification events. Breaking this number down per participant, this results in between 594 and 2,856 notification events per participant (M=1,436; SD=750; Md=1,240). As shown in Table 1, the server filtered 12,019 duplicate notification events and group summaries. 464 notifications were not annotated due to six participants sometimes missing the 48 hours annotation time window. This resulted in a total of 6,188 – or 93.02% – annotated unique notifications. Again, breaking this number down per participant, we saw that participants annotated between 253 and 761 notifications (M=476.00; SD=198.17; Md=464).

The annotated notifications were created by 94 different apps. The instant messaging app *WhatsApp* was used by all participants and dominated the number of created and annotated notifications. 65.22% of the annotated notifications were created by *WhatsApp*, followed by the *Google Play Store* (4.99%), and the instant messaging app *Telegram X* (3.80%).

Timings

Most annotated notifications were triggered around noon and in the afternoon. 25.26% were triggered between 10am and 2pm and 29.14% between 6pm and 9pm. The median time between a notification being triggered and finally being annotated was 17 hours and 53 minutes (see Figure 3). Two-thirds (72.66%) of the notifications were annotated within 24 hours. Participants annotated notifications throughout the day, with a third (32.97%) of the notifications being annotated between 5pm and 7pm.

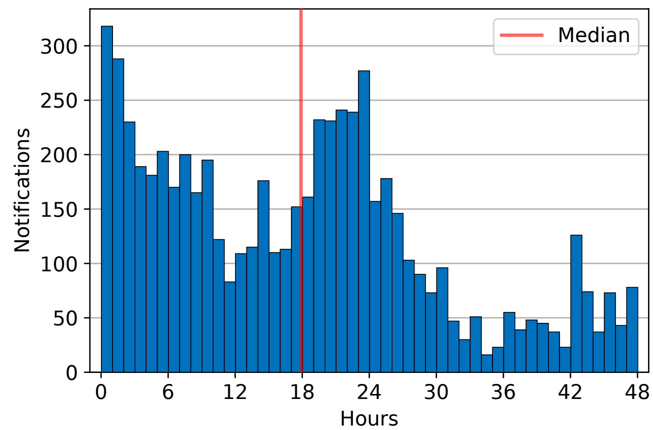

Figure 3: Histogram of the time delta between notifications being created on the participants' smartphones and finally being annotated.

Annotated Locations

Most of the notifications were annotated with the *Home* label (57.19%), followed by *Work* (14.87%), *On-the-go* (12.78%), *With friends* (9.2%), at a *Restaurant/cafe* (2.81%), *In public* (1.66%), and during *Sport* (1.12%). Only 0.37% of the notifications were annotated with the catch-all *Other* label. This indicates that the map shown next to the notifications in the annotation tool supported participants in assigning the notifications to a context.

Optional Comments

Eleven participants used the optional comment field to provide additional information for 4.12% of the annotated notifications. The comments mostly provided more details about the location (such as multiple labels applying) or mentioned that the location was off. This further indicates that participants were able to recall the context for a given notification.

Censored Content

Participants made use of the option to censor parts of the content for 32.01% of the annotated notifications. Only 0.82% of the notifications were censored completely. With 92.83%, most censored notifications were *WhatsApp* notifications.

We calculated how many notifications were censored per app. The five apps with the highest percentage of censored notifications were all communication apps: *WhatsApp* (45.56%), *Snapchat* (40.91%), *SMS* (37.50%), *Gmail* (34.55%), and *Facebook* (26.92%). Looking at which parts of the text were censored, we noticed that participants made use of the option to remove the names of their contacts in personal messages, with the message itself often left uncensored. This was interesting, as censoring the name and the entire message requires the same number of clicks in the annotation tool.

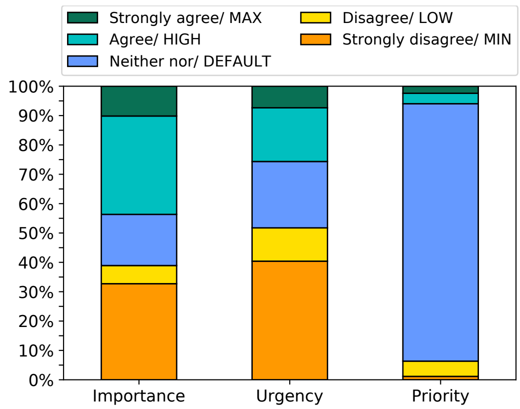

Figure 4: Distributions of the agreements that notifications are very important/ very urgent. The Android priority level set by apps for comparison.

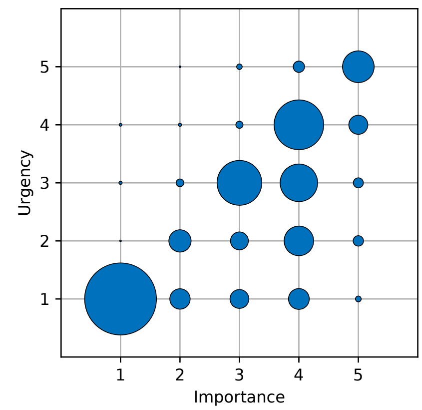

Figure 5: Correlation between the perceived importance and the perceived urgency of the annotated notifications. (1=strong disagreement; 5=strong agreement)

Participants seemed comfortable with sharing the messages as long as the senders' names were censored. This is an useful insight for future studies on mobile notifications.

Importance and Urgency Ratings

The distribution of the importance and urgency ratings can be seen in Figure 4. Looking at the agreements to the statement that a notification is *very important*, we found that participants (strongly) disagreed in 38.91% of the cases. 17.42% were rated neutral, and in 43.67% of the cases participants (strongly) agreed. Regarding the statement that a notification is *very urgent*, participants (strongly) disagreed in over half of the cases (51.75%). 22.6% were rated neutral, and only 25.65% of the annotated notifications received (strong) agreement ratings.

For comparison, we looked at the *priority* value that is set by apps for each notification. The five priority levels (`MIN`, `LOW`, `DEFAULT`, `HIGH`, `MAX`) can be compared to the 5-point Likert scale used for the importance and urgency ratings. We found that for most notifications (87.69%) the `DEFAULT` value was set. Thus, the priority level is not useful to decide on the actual importance or urgency of notifications.

Correlations

We calculated the Pearson correlation coefficient for the *importance* and *urgency* ratings, and the *priority* value. The results show that there is a strong positive correlation between the importance and the urgency of notifications ($r=0.82$; $p<0.001$). The correlation is visualized in Figure 5. Notably, participants perceived one-third of the notifications (32.60%) as neither important nor urgent (importance = urgency = 1). Only 6.33% of the notifications were regarded as both important and urgent (importance = urgency = 5).

There is neither a correlation between the rated importance and the priority ($r=0.18$; $p<0.01$) nor between the rated urgency and the priority ($r=0.16$; $p<0.01$).

Messaging Notifications

Of the 6,188 annotated notifications, 65.22% were created by the instant messaging app *WhatsApp*. All participants used *WhatsApp*, which reflects the dominant market share of the app in Germany. In the following, we provide a closer look at *WhatsApp* notifications according to four aspects. An overview of the ratings can be seen in Table 2.

(1) WhatsApp vs Other Apps. On average, *WhatsApp* notifications received higher importance and urgency ratings compared to notifications from other apps.

(2) Rich Media Messages. Apart from traditional text messages, *WhatsApp* allows users to send different rich media messages, including pictures, videos, and voice recordings. The notifications for these rich media messages contain corresponding emojis (📷, 🎥, 🎤) at specific positions of the text. This information allowed us to distinguish the notification types. The majority of notifications were for text messages (93.26%), followed by photos (4.44%), audio recordings (1.34%), and videos (0.97%). On average, notifications for audio recordings received the highest importance and urgency ratings, followed by text messages, photos, and videos.

(3) 1:1 vs Group Chats. *WhatsApp* allows conversations between two users (1:1 chats) and multiple users at once (group chats). Prior work used text heuristics to differentiate between 1:1 and group chats [36]. However, looking at the meta-data of *WhatsApp* notifications revealed that 1:1 chats are tagged with the string "`s.whatsapp.net`" and

Table 2: The average importance and urgency ratings for *WhatsApp* notifications based on specific features.

		Importance		Urgency		
	Feature	**M**	**SD**	**M**	**SD**	**N**
1	WhatsApp	3.10	1.34	2.56	1.34	4,036
	Other apps	2.30	1.48	2.13	1.36	2,152
2	Voice	4.41	.66	4.00	1.20	54
	Text	3.10	1.34	2.57	1.33	3,764
	Photos	2.86	1.36	2.12	1.17	179
	Videos	2.41	1.16	1.57	.72	39
3	1:1 chats	3.47	1.14	2.86	1.30	2,275
	Group chats	2.60	1.42	2.16	1.28	1,745
4	1:1 with name	4.22	.88	3.39	1.35	89
	1:1 without name	3.46	1.14	2.84	1.30	2,186
	Group w/ name	3.80	1.21	3.09	1.29	35
	Group w/o name	2.57	1.41	2.14	1.27	1,710
	Both w/ name	4.10	.99	3.31	1.34	124
	Both w/o name	3.07	1.34	2.54	1.33	3,896

group chats with "g.us". This allowed us to reliably differentiate between them, regardless of the user's device language. We found that more *WhatsApp* notifications were from 1:1 chats (56.37%), compared to group chats (43.23%). Only sixteen notifications (0.40%) were without a tag. On average, 1:1 chats were rated as more important and urgent than group chats (see Figure 7a), likely because users are not always addressed directly in group chats.

(4) Mentioning the User. The notification logging app automatically detected if notifications contain the first, last or nickname of the users and flagged notifications accordingly. Notifications that contain the user's name received higher importance and urgency ratings than the other notifications. This is not only true for *WhatsApp* 1:1 and group chats, but also over all apps (see Figure 7b).

Notification Clusters

Participants rated notifications from 94 different apps. We selected all apps whose notifications were rated by at least three participants. We then calculated the normalized importance and urgency ratings for the resulting 18 apps. For *WhatsApp*, we included the normalized overall rating and added 1:1 and group chats as well. The resulting 20 data points can be seen in Figure 6. For the Figure, we categorized and color-coded the apps. Applying *k-means* with a value of $k = 4$ revealed the notification clusters C1-C4.

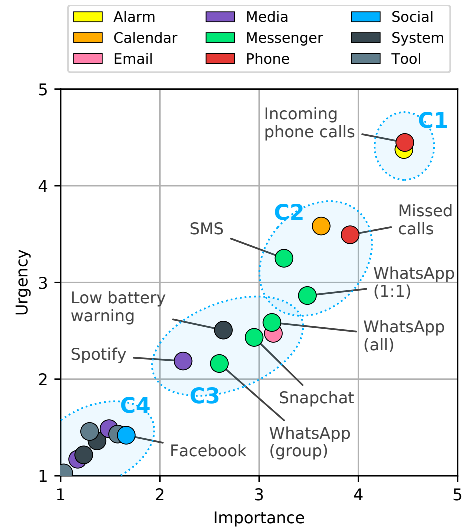

Figure 6: Normalized importance and urgency ratings of 18 apps that were rated by at least three participants. Additionally, we included *WhatsApp* 1:1 and group chats.

C1. *Incoming phone calls* received the highest ratings on average, closely followed by *alarms*. These types of notifications only contributed to 1.0% and 0.4% of the annotated notifications. It is easy to overlook these notifications when exploring the data set. However, they are of high importance and high urgency for the participants and often require their immediate attention.

C2. This cluster contains notifications about *missed phone calls*, calendar events (*Google Calendar*), *SMS* messages, and *WhatsApp* 1:1 messages. Notably, the *SMS* were not used for messaging. Instead, the notifications informed about calls going to the mailbox and phone plan updates. Notifications of this type are relevant to the user because the user is addressed directly, but they do not necessarily require the user's immediate attention.

C3. This cluster contains notifications from *WhatsApp* (overall), email (*Google Gmail*), *Snapchat*, *low battery warnings*, *WhatsApp* group chats, and *Spotify* music. Notifications of this type might not always be relevant to the user, and they are even less time-sensitive.

C4. The remaining eight apps were of the categories social, system, tool, and media. Notifications of this type are "nice-to-have" but neither of importance nor urgency. In some cases, they could even be considered annoying by users.

We found that notifications containing the name of the user (see Figure7b) or the name of a contact (see Figure7c) can be an indicator that the notification is of higher importance. The urgency ratings are affected similarly. As we have shown, this can be detected automatically by using the contacts stored on the device.

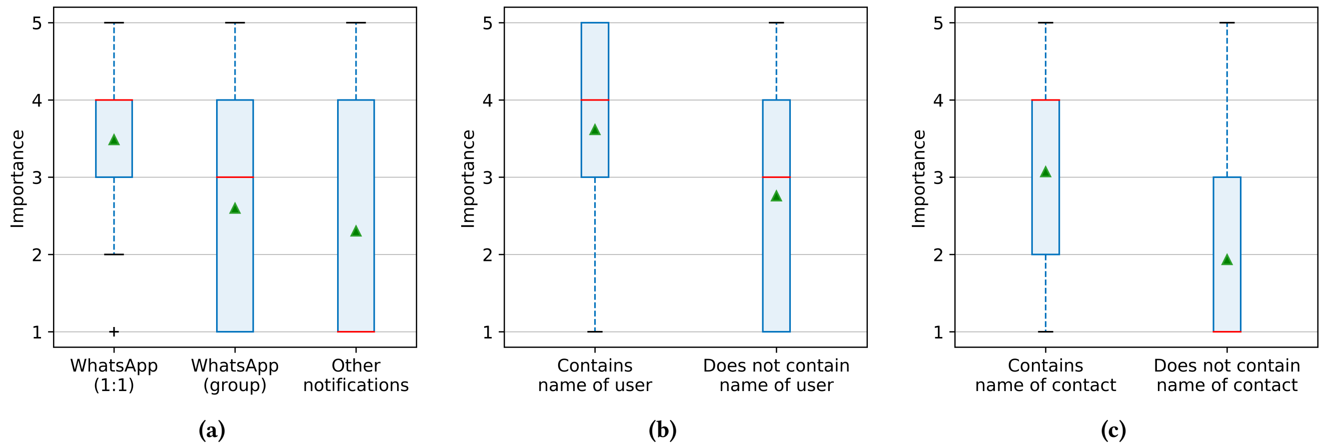

Figure 7: Differences in the importance ratings for the conditions (a) *WhatsApp* 1:1 and group notifications compared to notifications from other apps, (b) whether or not notifications contain the name of the user, and (c) whether or not notifications contain the name of a contact. (1=strong disagreement; 5=strong agreement)

Post-Study Questionnaire

After the participants annotated their notifications for a week, we sent out concluding emails. We thanked them for their participation and asked them to fill out a final post-study questionnaire. In the questionnaire, we asked them if they changed their smartphone usage behavior due to the participation in the study and if they were consistent in annotating their notifications. Eleven participants reported not changing their smartphone usage behavior during the study. One participant reported disabling specific notifications, and another participant mentioned uninstalling specific apps. Overall, participants were confident that they annotated their notifications consistently. Two participants mentioned sometimes having a hard time to rate the importance and urgency of notifications. Examples mentioned were notifications from music players about current songs and alarms. Participants also mentioned censoring the names of people to protect their privacy, something we were already able to see when looking at the annotated data.

In a final text field, participants informed us about their study experience. Annotating their notifications helped the participants to reflect on the notifications that they receive on a daily basis: *"It was interesting to see how many unimportant WhatsApp messages I receive throughout the day."* (P1)

Another participant realized during the study that she receives a large number of unimportant notifications that she subconsciously dismisses. *"Only after annotating I noticed how many notifications I receive that are not relevant to me at all (e.g., weather notifications). Notifications like that I simply dismissed without consciously reading them. [...] I would like to turn off such notifications in the future since I probably still perceive them subconsciously."* (P7)

Summary

We conducted an in-situ case study in which participants annotated their notifications for one week. We used the *Annotif* annotation system that allowed users to annotate notifications without interrupting them while preserving the notifications' content and context. We found a strong correlation between the perceived importance and urgency of notifications. Only a small percentage of the notifications were regarded as very important and very urgent. The instant messaging app *WhatsApp* dominated the number of created and annotated notifications. We saw differences in the perceived importance and urgency depending on the type of notification (e.g., 1:1 vs group and rich media messages). Notifications that contained the name of a contact or in which the user was addressed directly received higher importance and urgency ratings. Finally, we found four notification clusters that can be used to categorize notifications.

6 DISCUSSION

Using the *Annotif* system, we were able to collect 6,188 annotated unique notifications from 13 participants. This equals an annotation coverage of 93.02% for non-duplicate and non-group notifications, without triggering surveys throughout the day and potentially creating further interruptions. The annotation interface displayed the notifications' content and context (location and time). The comments provided by the participants in the optional free-text field indicated that the participants were able to reflect on the notifications well. Participants were able to screen the text of all logged notifications before sending them to us for analysis. Participants made active use of this functionality. However, we were positively surprised that participants rarely censored all text.

Instead, participants focused on preserving their contacts' privacy by censoring names. This enables a more throughout analysis of mobile notifications than simply relying on meta-data. Future user studies on mobile notifications might benefit from this finding. The *Annotif* system already detected contact names from the users' address books. This might be further extended in the future to automatically pre-censor notifications and, thus, reducing the number of interactions needed in the annotation tool.

The results of the importance and urgency ratings in the case study also pose interesting implications for smart notification management systems. As described in the *Notification Clusters* section, we found four notification clusters (C1-C4) in the data set. Critical notifications (C1) require the user's immediate attention. Examples include incoming phone calls and alarms. While critical notifications are of high importance and high urgency, they only contribute to a small fraction of the notifications users receive on a daily basis. Without a system that enables participants to assess *all* notifications, it is easy to overlook these notifications in larger data sets. On the other hand, we saw a large number of low priority notifications (C4). This is a long tail of apps that create unimportant and non-urgent notifications. Notifications of this kind may be considered nice-to-have or annoying by users.

In some cases, notifications may be lifted from one cluster to an adjacent cluster. We saw that mentioning the user or a contact increased the importance and urgency ratings. In reverse, notifications may drop to a lower priority level if they are received at the wrong time, e.g., personal notifications at work [42]. Finally, an app might display multiple types of notifications that are perceived differently by users. An example we saw in the study were notifications for rich media messages in *WhatsApp*, with voice recording notifications receiving higher ratings. This is a novel finding that was only uncovered by a notification data set with a high annotation coverage.

These clusters can aid designers of future smart notification management systems. We suggest that critical notifications (C1) should never be filtered or deferred by such systems. Low priority notifications (C4), however, can be deferred, shown in batches, or shown as summaries at the end of the day [3]. The biggest challenge for future smart notification management systems are high (C2) and medium (C3) priority notifications. These include communication-related notifications and are responsible for a large number of notifications users receive on a daily basis. Prior work has shown that there is a social pressure to respond as quickly as possible [8] and a fear of missing out [2]. Group chats might consist of highly relevant or completely irrelevant messages. Future work on messaging notifications can benefit from *Annotif*, as the system enables studying message contents in a privacy-respecting and unobtrusive manner.

7 LIMITATIONS AND FUTURE WORK

The annotation system as presented in this paper focused on the text of the notifications. However, the results of the case study indicate differences in the perception of rich media notifications. In the future, the annotation system could be improved by re-creating the notifications visually more similar to notifications on the smartphone. This includes displaying images associated with the notifications, such as profile pictures of contacts. Further, all participants in the case study were German. This was reflected in the apps used by the participants. The instant messaging app *WhatsApp* has a dominant market position in Germany. Other markets have different dominating messaging apps, e.g., *KakaoTalk* in South Korea and *WeChat* in China. Future studies should be conducted with a larger number and more diverse sets of participants over longer periods of time to create a more complete understanding of mobile notifications. Finally, future work should evaluate *Annotif* on a meta level. This includes the workload of annotating notifications over extended periods of time and the overall usability of the system.

8 CONCLUSION

Smartphone users receive a large number of notifications on a daily basis. These notifications can cause interruptions, which in return can have adverse effects on the user's productivity and well-being. Prior work investigated mobile notifications and means to reduce adverse effects, e.g., by silencing, filtering, or deferring notifications. However, researching mobile notifications is challenging, as studies are typically conducted in-situ and the content of notifications is of a private nature. In this paper, we introduced *Annotif*, a privacy-aware system for unobtrusively assessing mobile notifications in user studies. We reported the results of an in-situ case study in which participants annotated their notifications for one week. The results show that participants perceived 38.91% of their notifications as not important and over half (51.75%) as non-urgent. Only 6.33% of the notifications were rated as both very important and very urgent. We discussed influencing factors, including 1:1 and group messaging notifications, and implications for future smart notification management systems that continue to fulfill users' information need while respecting their digital well-being.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG) within project C04 of SFB/Transregio 161 and within the SimTech Cluster of Ex-cellence (EXC 310/2).

REFERENCES

[1] Piotr D. Adamczyk and Brian P. Bailey. 2004. If Not Now, when?: The Effects of Interruption at Different Moments Within Task Execution. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 271–278. https://doi.org/10.1145/985692.985727

[2] Julie H. Aranda and Safia Baig. 2018. Toward "JOMO": The Joy of Missing out and the Freedom of Disconnecting. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 19, 8 pages. https://doi.org/10.1145/3229434.3229468

[3] 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 (CHI EA '18)*. ACM, New York, NY, USA, Article LBW584, 6 pages. https://doi.org/10.1145/3170427.3188688

[4] Daniel Avrahami, Susan R. Fussell, and Scott E. Hudson. 2008. IM Waiting: Timing and Responsiveness in Semi-synchronous Communication. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 285–294. https://doi.org/10.1145/1460563.1460610

[5] Agathe Battestini, Vidya Setlur, and Timothy Sohn. 2010. A Large Scale Study of Text-messaging Use. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 229–238. https://doi.org/10.1145/1851600.1851638

[6] Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *ACM Comput. Surv.* 50, 6, Article 93 (Dec. 2017), 40 pages. https://doi.org/10.1145/3123988

[7] Jeremy Birnholtz, Jeff Hancock, Madeline Smith, and Lindsay Reynolds. 2012. Understanding Unavailability in a World of Constant Connection. *interactions* 19, 5 (Sept. 2012), 32–35. https://doi.org/10.1145/2334184.2334193

[8] Karen Church and Rodrigo de Oliveira. 2013. What's Up with WhatsApp?: Comparing Mobile Instant Messaging Behaviors with Traditional SMS. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 352–361. https://doi.org/10.1145/2493190.2493225

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

[10] Mary Czerwinski, Edward Cutrell, and Eric Horvitz. 2000. Instant messaging and interruption: Influence of task type on performance. In *OZCHI 2000 conference proceedings*, Vol. 356. 361–367.

[11] Mary Czerwinski, Ed Cutrell, and Eric Horvitz. 2000. Instant Messaging: Effects of Relevance and Timing. In *People and Computers XIV: Proceedings of HCI 2000* (people and computers xiv: proceedings of hci 2000 ed.), Vol. 2. 71–76. https://www.microsoft.com/en-us/research/publication/instant-messaging-effects-of-relevance-and-timing/

[12] Android Developers. 2018. NotificationListenerService. Retrieved August 22, 2019 from https://developer.android.com/reference/android/service/notification/NotificationListenerService.html

[13] Tilman Dingler and Martin Pielot. 2015. I'll Be There for You: Quantifying Attentiveness Towards Mobile Messaging. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 1–5. https://doi.org/10.1145/2785830.2785840

[14] Adrienne Porter Felt, Serge Egelman, and David Wagner. 2012. I've Got 99 Problems, but Vibration Ain't One: A Survey of Smartphone Users' Concerns. In *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM '12)*. ACM, New York, NY, USA, 33–44. https://doi.org/10.1145/2381934.2381943

[15] Joel E. Fischer, Chris Greenhalgh, and Steve Benford. 2011. Investigating Episodes of Mobile Phone Activity As Indicators of Opportune Moments to Deliver Notifications. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 181–190. https://doi.org/10.1145/2037373.2037402

[16] Joel E. Fischer, Nick Yee, Victoria Bellotti, Nathan Good, Steve Benford, and Chris Greenhalgh. 2010. Effects of Content and Time of Delivery on Receptivity to Mobile Interruptions. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 103–112. https://doi.org/10.1145/1851600.1851620

[17] Tony Gillie and Donald Broadbent. 1989. What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research* 50, 4 (01 Apr 1989), 243–250. https://doi.org/10.1007/BF00309260

[18] Sandy J. J. Gould, Duncan P. Brumby, and Anna L. Cox. 2013. What does it mean for an interruption to be relevant? An investigation of relevance as a memory effect. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 57, 1 (2013), 149–153. https://doi.org/10.1177/1541931213571034

[19] Rúben Gouveia and Evangelos Karapanos. 2013. Footprint Tracker: Supporting Diary Studies with Lifelogging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 2921–2930. https://doi.org/10.1145/2470654.2481405

[20] Joel M Hektner, Jennifer A Schmidt, and Mihaly Csikszentmihalyi. 2007. *Experience sampling method: Measuring the quality of everyday life.* Sage.

[21] Shamsi T. Iqbal and Brian P. Bailey. 2008. Effects of Intelligent Notification Management on Users and Their Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 93–102. https://doi.org/10.1145/1357054.1357070

[22] Daniel Kahneman, Alan B. Krueger, David A. Schkade, Norbert Schwarz, and Arthur A. Stone. 2004. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science* 306, 5702 (2004), 1776–1780. https://doi.org/10.1126/science.1103572 arXiv:http://science.sciencemag.org/content/306/5702/1776.full.pdf

[23] S Kalyanaraman, J Ivory, and L Maschmeyer. 2005. Interruptions and online information processing: The role of interruption type, interruption content, and interruption frequency. In *Proc. of 2005 Annual Meeting of International Communication Association*, Vol. 2005. 1–32.

[24] Kleomenis Katevas, Ioannis Arapakis, and Martin Pielot. 2018. Typical Phone Use Habits: Intense Use Does Not Predict Negative Well-being. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 11, 13 pages. https://doi.org/10.1145/3229434.3229441

[25] Kostadin Kushlev, Jason Proulx, and Elizabeth W. Dunn. 2016. "Silence Your Phones": Smartphone Notifications Increase Inattention and Hyperactivity Symptoms. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1011–1020. https://doi.org/10.1145/2858036.2858359

[26] Luis Leiva, Matthias Böhmer, Sven Gehring, and Antonio Krüger. 2012. Back to the App: The Costs of Mobile Application Interruptions. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '12)*. ACM, New York, NY, USA, 291–294. https://doi.org/10.1145/2371574.2371617

[27] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. 2016. PrefMiner: Mining User's Preferences for Intelligent Mobile Notification Management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 1223–1234. https://doi.org/10.1145/2971648.2971747

[28] Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. 2015. Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 813–824. https://doi.org/10.1145/2750858.2807544

[29] Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. 2016. My Phone and Me: Understanding People's Receptivity to Mobile Notifications. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1021–1032. https://doi.org/10.1145/2858036.2858566

[30] Tadashi Okoshi, Julian Ramos, Hiroki Nozaki, Jin Nakazawa, Anind K. Dey, and Hideyuki Tokuda. 2015. Reducing Users' Perceived Mental Effort Due to Interruptive Notifications in Multi-device Mobile Environments. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 475–486. https://doi.org/10.1145/2750858.2807517

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

[32] Veljko Pejovic and Mirco Musolesi. 2014. InterruptMe: Designing Intelligent Prompting Mechanisms for Pervasive Applications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 897–908. https://doi.org/10.1145/2632048.2632062

[33] Martin Pielot, Karen Church, and Rodrigo de Oliveira. 2014. An In-situ Study of Mobile Phone Notifications. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 233–242. https://doi.org/10.1145/2628363.2628364

[34] Martin Pielot, Rodrigo de Oliveira, Haewoon Kwak, and Nuria Oliver. 2014. Didn't You See My Message?: Predicting Attentiveness to Mobile Instant Messages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3319–3328. https://doi.org/10.1145/2556288.2556973

[35] Martin Pielot and Luz Rello. 2017. Productive, Anxious, Lonely: 24 Hours Without Push Notifications. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, Article 11, 11 pages. https://doi.org/10.1145/3098279.3098526

[36] Martin Pielot, Amalia Vradi, and Souneil Park. 2018. Dismissed! A Detailed Exploration of How Mobile Phone Users Handle Push Notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 3, 11 pages. https://doi.org/10.1145/3229434.3229445

[37] 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 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3055–3064. https://doi.org/10.1145/2556288.2557189

[38] Cary Stothart, Ainsley Mitchum, and Courtney Yehnert. 2015. The attentional cost of receiving a cell phone notification. *Journal of experimental psychology: human perception and performance* 41, 4 (2015), 893.

[39] Jan Willem Streefkerk, D Scott McCrickard, Myra P van Esch-Bussemakers, and Mark A Neerincx. 2012. Balancing awareness and interruption in mobile patrol using context-aware notification. *International Journal of Mobile Human Computer Interaction (IJMHCI)* 4, 3 (2012), 1–27.

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

[41] Dominik Weber, Alireza Sahami Shirazi, and Niels Henze. 2015. Towards Smart Notifications Using Research in the Large. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '15)*. ACM, New York, NY, USA, 1117–1122. https://doi.org/10.1145/2786567.2794334

[42] Dominik Weber, Alexandra Voit, Jonas Auda, Stefan Schneegass, and Niels Henze. 2018. Snooze! Investigating the User-defined Deferral of Mobile Notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 2, 13 pages. https://doi.org/10.1145/3229434.3229436

[43] 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 (UbiComp '18)*. ACM, New York, NY, USA, 1271–1278. https://doi.org/10.1145/3267305.3274118

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

[45] Dominik Weber, Alexandra Voit, Philipp Kratzer, and Niels Henze. 2016. In-situ Investigation of Notifications in Multi-device Environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 1259–1264. https://doi.org/10.1145/2971648.2971732

[46] 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 (MobileHCI '16)*. ACM, New York, NY, USA, 936–941. https://doi.org/10.1145/2957265.2962660

[47] WhatsApp. 2017. Connecting One Billion Users Every Day. Retrieved August 22, 2019 from https://blog.whatsapp.com/10000631/Connecting-One-Billion-Users-Every-Day

[48] SungHyuk Yoon, Sang-su Lee, Jae-myung Lee, and KunPyo Lee. 2014. Understanding Notification Stress of Smartphone Messenger App. In *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 1735–1740. https://doi.org/10.1145/2559206.2581167