My Phone and Me: Understanding User's Receptivity to Mobile Notifications

ABSTRACT

Notifications are extremely beneficial to the users, but they often demand their attention at inappropriate moments. In this paper we present an *in-situ* study of mobile interruptibility focusing on the effect of cognitive and physical factors the response time and the disruption perceived from a notification. Through a mixed method of automated smartphone logging and experience sampling we collected 10372 *in-the-wild* notifications, 474 questionnaire responses on notification perception from 20 users. We found that the presentation of a notification, its alert type, sender-recipient relationship as well the type, completion level and complexity of a task the user is engaged in influence the response time and the disruption perceived from a notification. We found that even a notification that contains important or useful content can cause disruption. Finally, we observe a substantial role of the psychological traits of the individuals on the response time and the disruption perceived from a notification.

Author Keywords

Mobile Sensing; Notifications, Interruptibility,
Context-aware Computing.

ACM Classification Keywords

H.1.2. Models and Principles: User/Machine Systems; H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

INTRODUCTION

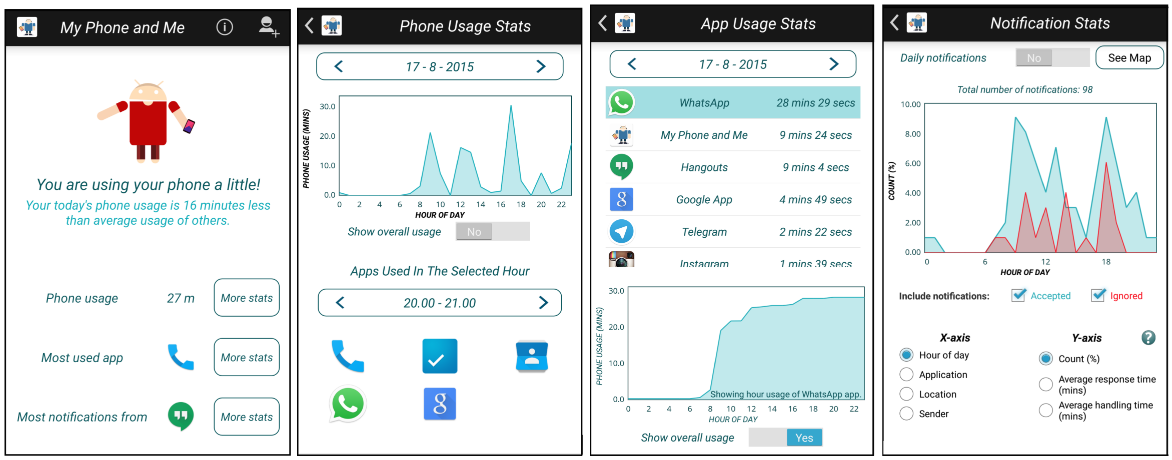
Smartphones enable a new form of effortless information awareness. Throughout the day, a smartphone user receives a variety of information, from email messages, over social network event invites to birthday reminders. Notifications are at the core of this information awareness, as they use audio, visual and haptic signals to steer the user's attention towards the newly-arrived information.

Notifications are extremely beneficial to the users: however, at the same time, they are a cause of potential disruption, since they often require users' attention in inopportune moments. Indeed, previous studies have found that interruptions at inopportune moments can adversely affect task completion time [12, 13, 27], lead to high task error rate [9], and impact the emotional and affective state of the user [6, 8]. Also, users might get annoyed when they receive notifications presenting information that is not useful or relevant to them in the current context [14]. At the same time, studies have shown that users cannot keep their smartphones away for a long time, because the users start feeling stressed and anxious about missing important information until they finally pick up the phone to check for any new notifications [29]. This tension is exacerbated by the fact that individuals deal with hundreds of notifications in a day, some of which are disruptive [25]

Previous studies have investigated the user's *receptivity* to mobile notifications [16, 34, 31]. As defined by Fischer [16], receptivity encompasses a user's reaction to an interruption and her subjective experience of it. For instance, a user might quickly respond to a notification when she is idle, but she can still get annoyed because of the content of the notification. Previous studies have shown that the user's receptivity to a notification is determined by: (i) how interesting, entertaining, relevant and actionable its content is for the user [16]; (ii) the type of application that triggers it – communication applications are considered as the most important [34]; (iii) time criticality and social pressure [31]. At the same time, some studies have proposed various mechanisms to infer *opportune moments*, i.e., moments in which a user quickly and/or favorably reacts to a notification [15, 25, 30, 32]. In order to infer interruptibility these studies have used machine learning classifiers provided with different contextual factors including user's transitions between activities [20], engagement with a mobile device [15], time of day, location and activity [30], and notification content [25].

However, none of these studies have deployed the proposed mechanisms in the real world scenario of in-the-wild notifications of a regularly used application. The key reason behind this is the fact the accuracy of these mechanisms is still lower than the user's expectations. In the real world scenario, the users would not accept a system that might defer or stop an important notification. Previous studies have shown that users are willing to tolerate some interruption, so that they do not miss any important information [22]. We believe that interruptibility management systems fail to achieve a very high accuracy in predicting the opportune moment because *there is still a lack of understanding concerning the factors influencing the user's receptivity to mobile notifications in different physical and cognitive situations*.

In order to bridge this gap, in this work we conduct an in-situ study to collect objective and subjective data about mobile notifications. We design and develop *My Phone and Me* (Figure 1), an application that uses a novel experience sampling method (ESM) approach to uncover the factors and motivations impacting the user's handing and her sentiment towards



(a) (b) (c) (d)

Figure 1. My Phone and Me application: (a) main screen, (b) phone usage statistics, (c) application usage statistics, (d) daily notifications.

a notification. Through *My Phone and Me*, we collect 19494 notifications, 611 responses for the questionnaire and 11 personality test results, from 74 users in around two months long time period. Using the subset of this data corresponding to users who completed to at least 14 ESM questionnaires, we investigate users' interaction with mobile notifications in different physical and cognitive contexts. More specifically, the key contributions of this work are the investigation of:

- the impact of a notification's alert modality on the user's ability to perceive a notification alert;
- the impact of the alert modality, sender-recipient relationship, presentation of a notification, an ongoing task type, completion level and task complexity on the response time;
- the impact of the sender-recipient relationship, and an ongoing task's type, completion level and complexity on the perceived disruption;
- the role of the sender-recipient relationship, notification content and the perceived disruption on the user's decision to accept or dismiss a notification;
- the impact of the user's personality on the perceived disruption and response time of a notification.

The findings of our study are wide-ranging, and may have a direct impact on the way future notification management mechanisms are constructed. First, we observe that a sender-recipient relationship, notification priority, and an ongoing task's type and complexity influence the response time of the notification, but there is no impact of the ongoing task's completion level on the response time. Moreover, our results show that the recipient's relationship with the sender of a notification, the ongoing task's type, completion level and complexity influences the perceived disruption. Our findings imply that the higher the level of disruption perceived by the user the higher the probability of the notification being dismissed. From our results, we also observe that nevertheless users tend to click highly disruptive notifications if they contain valuable information. While users are aware of notifications even when their phone is in silent mode, our analysis shows that the alert modality has a significant impact on the time taken by the users to view the notification. Finally, we observe a substantial role of the psychological traits on how a person reacts to a mobile notification, calling for highly personalized interaction between a smartphone and its user.

REASONING ABOUT USER'S RECEPTIVITY TO MOBILE NOTIFICATIONS

An interruption tries to steer a user from an ongoing primary task to the secondary task signaled by it [9]. As suggested by Clark [11], users can respond to an interruption in four possible ways: 1. handle it immediately; 2. acknowledge it and agree to handle it later; 3. decline it (explicitly refusing to handle it); 4. withdraw it (implicitly refusing to handle it).

A user can respond to mobile notifications in a fairly different way as compared to an in-person interruption. For communication-related interruptions, for example, users might perceive more disruption from an in-person interruption than from a mobile notifications because of the presence of an interrupter in the former case. Mobile notifications enable flexibility in the way an interruption is handled because of the lack of the physical presence of the sender, and the asynchronous nature of mobile messaging communication[1]. Thus, the exact moment of handling an interruption can be negotiated and the recipient can decide when and how to attend a notification.

However, this flexibility introduces many other issues. First, notifications can go unnoticed in case a user does not register an alert. Second, usually non-persistent, notifications may be forgotten about – a user riding a bicycle, might decide to attend to a notification once he arrives to the destination, yet eventually forgets to do so. Finally, although designed to signal an interruption, but not interrupt themselves, mobile notifications can still induce unnecessary disruption to a user's

[1]Certain social norms and expectations from the sender side, however, constraint the flexibility that the receiver has in reacting to a message [32]

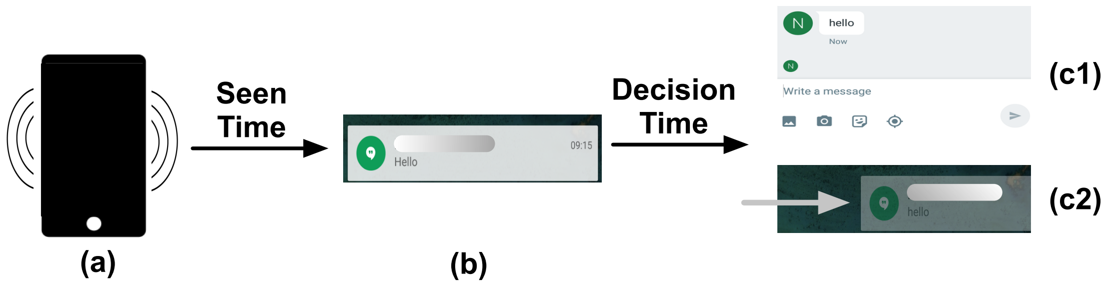


Figure 2. The three time measurements of a notification captured by the My Phone and Me application. The time of notification arrival (a), the time when a notification is seen (b), and the time when the user accepted ($c1$) or dismissed ($c2$) a notification. The time difference between (a) and (b) is seen time and the time difference between (b) and ($c1$ or $c2$) is the decision time.

routine. For instance, the disruption can happen when a user decides to attend a notification immediately, despite being in the middle of another task, only to find that the notification is about an unrelated promotional offer. Moreover, a disruption may happen even if a notification is not attended, as a thought of a lingering notification may interfere with the user's current task performance [35].

In this study, we investigate the factors influencing the user's response to a mobile notification, where the response is defined by the time taken to register and react to a notification, and the way in which the notification is handled (i.e., clicked or dismissed). Moreover, we investigate the user's motivation for being self-disruptive by clicking the disruptive notifications.

Our assumption is that the response time of a notification and the disruption perceived by the user are influenced by the different aspects of the notification as well as the user's context. To capture this measure we develop an Android experience sampling method (ESM) application that monitors the actual notifications a user receives on her phone, records a user's reaction to notifications, and then queries the user to identify her rating of the disruption caused by the notification. We augment this with the questions about the motivation for handling a notification in a particular way. Further, our ESM questionnaires ask the user to provide the data on the type, complexity and the completion level of the ongoing task, and the user's relationship with the sender. Finally, we collect participants' personality trait measures at the end of the experiment.

First, we investigate the ability of users to adjust their response times to a notification, and see how quickly they can triage different notifications in different situations. As shown in Figure 2, we take three time measurements for each notification: the time of notification arrival (a), the time when a notification is seen (b), and the time when the user accepted (c1) or dismissed (c2) a notification. Note that in order to detect the moment at which a notification is seen, we use the unlock event of the phone and assume that all newly available notifications in the notification bar are seen when the user unlocks the phone. In case a notification arrives when the user is already using the phone (i.e., the phone is unlocked), the seen time of this notification would be computed as zero. We term the time from the notification arrival till the moment the notification was reacted upon as *the response time* for the notification. For our analysis, we break the response time into two intervals:

- Seen time (ST) – time from the notification arrival till the time the notification was seen by the user.
- Decision time (DT) – time from the moment a user saw a notification till the time she reacted upon it.

We examine the way interruption timing, with respect to the primary task, determines the user's response to the notification. Moreover, we are interested in the way users triage disruptive notifications. Can users quickly discern about notifications that are disruptive? We hypothesize that humans might still attend a notification, even if they know that the primary task is going to be disrupted. For example, in their study of WhatsApp notifications Pielot et al [32] show how, due to an inner pressure raised by social expectations, users quickly respond to instant messaging (IM) communication, or frequently check their phones, inducing self interruptions just in order to satisfy the social expectations. In our work, we are looking beyond just IM notifications and investigate the way any disruptive message is handled. Through our ESM study we identify the motivation behind reacting to a disruptive notification, and the reasoning and the external factors that lead to the exact reaction.

We aim for a comprehensive investigation of interruptibility from a user perspective, thus compare the effect of different aspects of a notification on its response time and disruptiveness. Finally, we investigate the potential role of individual psychological traits on how users perceive and react to disruptive notifications.

DATA COLLECTION

In order to investigate the nature of disruptive notifications and factors that determine the user's receptivity to mobile notifications in different physical and cognitive situations, we conducted an in-situ field study. More specifically, we developed an Android app called *My Phone and Me* – an Android experience sampling method (ESM) application that collects information about *in-the-wild* notifications, users' interaction with them in natural situations (while they are performing

Group	Features
Time	Arrival, seen and the removal time of a notification.
Notification response	Whether the notification was clicked or dismissed.
Notification details	Sender application and the title of a notification.
Alert type	Signals used by a notification to alert the user: sound, vibrate, and flashing LED.
Context data	Physical activity, location, presence of surrounding sound, WiFi connectivity, Proximity to the phone, surrounding light intensity. This data is collected on arrival and removal of a notification from the notification bar.

Table 1. Description of features from the My Phone Me dataset.

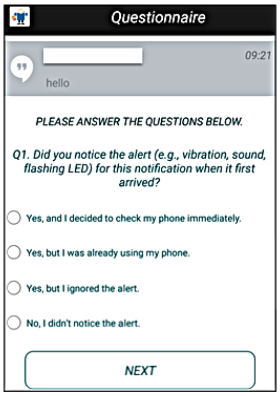


Figure 3. The screenshot of a questionnaire triggered by the My Phone and Me application.

Question	Options
Did you notice the alert (e.g., vibration, sound, flashing LED) for this notification when it first arrived?	(i) Yes, and I decided to check my phone immediately. (ii) Yes, but I was already using my phone. (iii) Yes, but I ignored the alert. (iv) No, I didn't notice the alert.
How did you handle the notification when you first saw it?	(i) I decided to immediately click it. (ii) I decided to dismiss it because it didn't require any further action. (iii) I decided to dismiss it because it was not relevant or useful. (iv) I decided to return to it later. (v) Other (descriptive).
Select all factors that made you decide to click/dismiss the notification.	(i) The sender is important. (ii) The content is important. (iii) The content is urgent. (iv) The content is useful. (v) I was waiting for this notification. (vi) The action demanded by the sender does not require a lot of effort. (vii) At this moment, I was free. (viii) Other (descriptive).
What best describes your relationship to the sender?	(i) Partner (ii) Immediate family (children, parents) (iii) Extended family (nieces/nephews, cousins, aunts/uncles) (iv) Friend (v) Acquaintance (vi) Superior at work (vii) Colleague (viii) Subordinate at work (ix) Client (x) Service provider (xi) Sender is not a person (xii) Other relationship (descriptive).
Please describe what the notification was about.	Descriptive response.
Please describe what activity you were involved with when you received the notification.	Descriptive response.
When the notification arrived, I was:	(i) Starting a new task/activity. (ii) In the middle of a task/activity. (iii) Finishing a task/activity. (iv) Not doing anything.
The task/activity I was doing when the notification arrived was complex.	Five-level Likert scale rating between "strongly disagree" and "strongly agree".
I found the notification disruptive.	Five-level Likert scale rating between "strongly disagree" and "strongly agree".

Table 2. Questions and their options from My Phone and Me questionnaire.

their day-to-day activities), and the physical and cognitive context details.

The My Phone and Me application uses Android's Notification Listener Service [1] to access notifications, and Google's Activity Recognition API [3] and ESSensorManager [24] to obtain the context information. Table 1 lists the groups of features captured by the application. Note that in order to infer the user's response to a notification, the My Phone and Me application checks whether the application that triggered the notification was launched after the removal time of that notification. We are aware that some notifications are dismissed because they do not require any further action. For this reason, we capture seen time and use the difference between seen time and removal time to understand how long does it take to the user for reading and dismissing a notification.

To collect subjective data from the users, the My Phone and Me application triggers four questionnaires in a day. A questionnaire is triggered only when a notification is handled; it contains questions about why the notification was clicked or dismissed by presenting a screenshot of that notification (see Figure 3). The application triggers a questionnaire for a randomly selected notification in every four hours time window between 8.00 am and 8.00 pm and the last questionnaire at a random time between 8.00 pm and 10.00 pm. The application did not trigger any questionnaire after 10pm so that the participants do not feel annoyed for responding the surveys late at night. The application automatically used the local time zones because it relies of the phone's time. Moreover, if the user is busy, the questionnaire can be dismissed by simply swiping it from the notification bar and no questionnaire is shown to the user for the next 30 minutes.

A questionnaire comprises of seven multiple-choice and two free-response questions. The list of questions and their options are shown in Table 2. Since we ask the users to enter the free form text for two questions, it could increase time to respond to a questionnaire and may become a source of annoyance. Therefore, the application allows the users to dictate the responses to these questions. The application relies on the Android's SpeechRecognizer API [2].

Additionally, the My Phone and Me application asks the users to take a personality test based on the 50 item-based Big-Five Factor Markers from the International Personality Item Pool, developed by Goldberg [18, 5]. A notification to take this test is triggered once the user has responded to 28 questionnaires. A user can also take the test at any time by clicking on the personality test button present in the application's action bar.

Recruitment of the Participants

The My Phone and Me application was published on the Google Play Store from 12th August 2015. It was installed by

74 participants without any monetary incentive. These participants come from both sexes, with the age span between 19 and 50 years. As shown in Figure 1, the My Phone and Me tells the users about their addiction to the phone. It allows users to check statistics on their phone usage and interruptions. The application visualizes a user's phone activities based on different criteria, such as her hourly phone usage (Figure 1 C), hourly usage of individual applications (Figure 1 D) and how much she interacts with notifications (Figure 1 B). We believe that displaying this information has a minimal interference on users' actual behavior for interacting with notifications, but provides a valuable functionality in order to make them keep the application installed on their phones.

In order to ensure the privacy compliance, the My Phone and Me application goes through a two-level user agreement for access the user's notifications. Firstly, the user has go give explicit permission as required by the Android operating system, and secondly, the application shows a list of information that is collected and asks for the user consent. Moreover, we show the original content of a notification to the user along with the questionnaire in order to avoid any recall bias in the data, but we do not collect the notification content for privacy reasons.

DATASET

The data collection was carried out for for around two months, during which we collected 19494 notifications, 611 responses for the questionnaire (comprises of a set of nine questions listed in Table 2) and 11 personality test results (50 item-based Big-Five Factor Markers by Goldberg [18, 5]) from 74 users who installed the My Phone and Me application. Many users stopped responding to the questionnaires after a few days and some did not respond at all. Therefore, we select a subset of the data for the analysis and include data of users who responded to at least 14 questionnaires. There are 20 users who satisfied this constraint. So, our final dataset comprises of 10372 notifications, 474 questionnaire responses and 11 personality test results. Additionally, during the setup phase we asked participants to enter their age and gender: in our dataset there are 11 male and 9 females, and are aged between 19 to 50 year old.

As we are primarily using the questionnaire responses, we compared the click rate (i.e., percentage of notifications that are clicked out of total notifications) of the overall notifications with the notifications that were linked to questionnaires. The click rate for overall notifications is 62.52%, and for notifications linked with questionnaires is 70.04%.

We are aware that our dataset has potential limitations that stem from the inherent nature of an *in-the-wild* study. The dataset remains unbalanced because it is not possible to obtain equal number of questionnaire responses from all users for each category of a test. For instance, there is a very small chance that our application triggers a questionnaire for each type of a sender from the recipients social circle. Furthermore, in practice, a user might not even receive notifications from each of the sender types during the period of the study. Therefore, we use the data from 20 users who responded to at least 14 questionnaires, i.e., the minimum number of questionnaire that were answered by a user in this set. We also make an hypothesis of data independence, which might not hold in reality and can be tested only in a controlled setting.

UNDERSTANDING RESPONSE TIME

In this section we present the results of our investigation on the role of alert modality, sender type (i.e., sender – recipient relationship such as family and friends), and type, complexity and completion level of the ongoing task, on the seen and decision time of a notification.

The Role of Alert Modality in Perceiving a Notification Alert

A notification can alert the user by the means of vibration, sound, and/or flashing LED. In order to investigate how users perceive alerts with different the alert modalities, we used the responses provided by the users for Q1 (*Did you notice the alert (e.g., vibration, sound, flashing LED) for this notification when it first arrived?*). According to our dataset, when the notification (referred in the questionnaires) were triggered the user's phone was for 25.54% of the times in the silent mode, 21.50% vibrate mode, 41.94% sound mode and 11.03% sound with vibrate.

Users reported that they missed notification alerts for 14.63%, 15.38%, 23.75%, 21.05% of times their phone was in silent, vibrate, sound, and sound with vibrate mode respectively. This provides an evidence that when the phone is in silent mode, users are still aware of the notification alerts.

What Factors Influence the Seen Time?

We investigate the role of alert modality, sender, and type, complexity and completion level of the ongoing task, in influencing the seen time of a notification.

The Impact of Alert Modality on Seen Time. In order to perform this analysis, from our dataset of 10372 notifications we use all the notifications of the 20 users who responded to at least 14 questionnaires. However, we ignore the notifications that arrived when the user was already engaged with the phone because we could not calculate the seen time of these notifications. This leaves us with 4929 notifications. A one-way Analysis of Variance (ANOVA) of the seen time was carried out for each alert modalities. The results show that the alert modality has an impact on the seen time of notifications, with $F(3, 4925) = 26.41$, $p < 0.001$. A Tukey post-hoc test (by setting the α as 0.05) revealed that the seen time is statistically significantly higher for silent notifications (average 7.332 mins). The seen time for the notifications alerting with vibrate only mode is the lowest (average 3 minutes and 21 second). Sound only and sound with vibrate notifications are the second (average 5 minutes and 57 seconds) and third (average 4 minutes and 50 seconds) most quickly seen by users. Quite interestingly, a recent 15-user study by Pielot et al. also found that notifications tend to be seen faster when the phone is in the vibrate mode [31]. Here, we confirm this finding, but also point to the above missed notification percentage in the silent mode (14.63%) and show that setting the phone to silent does not help escape the interruptions.

The Impact of Ongoing Task Type on Seen Time. To investigate the impact of the ongoing task on the notification's seen time, we require the type of task that the users were involved with when the notification arrived. We classified the information users provided through ESM questionnaires about the ongoing task into the following six categories: work, communication, traveling, maintenance/personal, leisure, idle. The classifications was done manually, by two coders who initially disagreed on five entries. Two common labels for these were found after a discussion with other three coders. Note that our app allows users to skip the step of providing the information on the question about their current task by selecting "Prefer not to say" option. In such cases, we discard the entry from our analysis of the effect of the ongoing task on interruptibility.

A one-way ANOVA of the seen time is carried out for each task types. The results show that the ongoing task type has an impact on the seen time of notifications, with $F(5, 217) = 2.963$, $p = 0.013$. A Tukey post-hoc test (by setting the α as 0.05) reveals that the seen time is the lowest when the notifications arrive while the user is communicating (average 47 seconds) and highest while the user is idle (average 9 minutes and 30 seconds). Other task types do not have a statistically significant effect on the seen time of notifications and have an average seen time of 5 minutes 45 seconds. As shown in a recent study [33] notifications are more welcoming when the recipient are bored. However, our results show that while the users might be willing to accept more notifications when idle, the time needed to attend such notifications might be higher compared to the time needed to attend a notification while a user is busy.

The Impact of Ongoing Task Complexity on Seen Time. To analyze the effect of an ongoing task complexity, we first encode the reported task complexity, which was reported as value on the Likert scale (*Strongly disagree*=1, *Somewhat disagree*=2, *Neutral*=3, *Somewhat agree*=4 and *Strongly agree*=5) to the question "The task/activity I was doing when the notification arrived was complex". The Pearson correlation coefficient is computed to evaluate the relationship between the complexity of an ongoing task and the seen time of a notification. The results depict that there is a weak, negative correlation between the two variables, $r = -0.189$, $p = 0.005$. Thus, the increase in the seen time of notifications is correlated with the decrease in rating of ongoing task's complexity. We believe that this correlation exists because the users become more concentrated while performing a complex task and thus, quickly perceive the interruptions. On the other hand, when the user are not performing any complex task, they become less attentive to the interruptions. Finally, we found that factors such as the completion level of the ongoing task and the sender type do not have a statistically significant effect on the seen time of notifications.

The Impact on Decision Time

We analyze the effect of the type, complexity and completion level of the ongoing task, and the sender type on the time a user takes to decide on how to react to a notification. We find that all of these factors do not have a statistically significant effect on the decision time of notifications with the exception of the sender.

A one-way repeated measures ANOVA of the decision time was carried out for each sender types. The results show that the sender type has an impact on the seen time of notifications, with $F(10, 212) = 2.429$, $p = 0.00936$. A Tukey post-hoc test (by setting the α as 0.05) revealed that out of the 11 sender types (shown in Table 2), notifications from partner lead to the fastest decision time (mean DT is 3.315s), followed by immediate family members with an average decision time of 4.891 seconds. On the hand, notifications from extended family members and service providers have the highest decision time, 11.93 and 8.146 seconds respectively. There was no statistically significant difference in the decision time of the notifications from other senders. These results demonstrate that notifications are quickly handled when they are sent by the close relatives of the user. In other cases users take more time in reading the content before deciding how to handle it. We hypothesize that this behavior stems from the content of notifications from close friends or family members, which might be more predictable, and a part of a daily routine (e.g., "pick kids from school"). Whereas, the users have to spend more time on the notifications from less frequently contacted sources, as the content may be less familiar to them.

The Role of Notification Presentation

In our dataset, 2953 (out of 7795) notifications were received when the user was engaged with the phone. Out of these 2953 notifications, 860 are so-called "low-priority" while 2093 are "high-priority" notifications [4]. Here, a high-priority notification is a foreground notification that gets in the way of the user's ongoing activity and the user cannot perform any action to get it out of the way without clicking or dismissing it (e.g. Viber messages). A low-priority one simply appears on the notification bar without getting in the way of the user's ongoing activity (e.g. Gmail notifications).

We investigate the effect of the notification presentation on the response time of a notification. The result of a two sample t-test shows that there is a statistically significant effect of notification priority on the response time, $t(2951) = 17.694$, $p < 0.001$, with high-priority notifications getting quicker response than low-priority notifications. The mean response time for high-priority notifications is 11.94s versus 25.91s for low-priority notifications.

WHY A NOTIFICATION BECOMES DISRUPTIVE?

In order to analyze the effect of different factors on the perceived disruption, we first encode the perceived disruption, which is reported as an answer to "I found the notification disruptive", and takes one of the following possible values from the Likert scale: *Strongly disagree*=1, *Somewhat disagree*=2, *Neutral*=3, *Somewhat agree*=4 and *Strongly agree*=5.

The Role of Ongoing Task Complexity. We investigate whether the complexity of an ongoing task is associated with the perceived disruption reported by the users. A Kendall's Tau correlation coefficient was computed to assess the relationship between the ongoing task complexity and perceived disruption. We found a strong, positive correlation between the two variables, $R\tau = 0.477$, $p < 0.001$. This demonstrate that the users are more likely to get more disrupted from a notification that arrives when they are engaged in an intricate task, and less disrupted when they are performing a simple task.

The Role of Ongoing Task Completion Level. A one-way ANOVA of the reported disruption was carried out for each class of task completion level (starting, in the middle, finishing, and not doing anything). The results show that the completion level of an ongoing task has a significant impact on the disruption perceived by the users from the notifications, $F(3, 451) = 19.43$, $p < 0.001$. A Tukey post-hoc test (by setting α as 0.05) reveals that the perceived disruption is the highest when the user is currently involved in a task. The perceived disruption is the lowest when the user is starting a task or idle, and there is no statistically significant difference between these groups. These results show that the perceived disruption when the user is highly engaged in a task is very high not only from the desktop notification, as discussed for example in [14, 26], but also from the mobile notifications.

The Role of Sender. We compute a one-way ANOVA of the reported disruption for each type of sender (see Table 2). According to the results, $F(10, 444) = 3.987$, $p < 0.001$, the type of sender has a significant impact on the disruption perceived by the users from the notifications. A Tukey post-hoc test (by setting the α as 0.05) reveals that the perceived disruption is the highest when the sender is not a person and a subordinate at work (no statistically significant difference between these two groups) and the lowest when the sender is an extended family member. Moreover, colleagues and service providers are the second most disruptive sender groups. There is no significant difference between the other groups. Previous studies showed that users express a negative *sentiment* towards messages not coming from their family and friends [16], and that more "distant" the sender is, less likely it is that a notification will be *clicked on* [25]. Results from our study complement this with the finding that the perceived *disruption* varies with the sender of a notification.

The Role of Ongoing Task Type. A one-way ANOVA of the reported disruption is carried out for each type of ongoing task (see Table 2). The results show that the type of task that the user is engaged with on the arrival of a notification has a significant impact on the disruption the user perceives when the notification arrives, $F(5, 380) = 13.03$, $p < 0.001$. A Tukey post-hoc test (by setting α as 0.05) revealed that the perceived disruption is the highest when the user is working and the lowest while the user is idle. After work, traveling and then leisure are the tasks when the users perceive high disruption. When the users are not idle, they perceive least disruption while communicating and doing a personal or maintenance task. Since the communication can involve notifications themselves, e.g. two mobile users exchanging WhatsApp messages, the above result is not surprising. As shown in a recent study [33] that the user's are receptive to information when they are bored, our results go inline with these finding to show that perceived disruption is lowest when the user is idle.

Option	Count (%)
Sender is important	31.546
The content is important	27.129
The content is urgent	14.511
The content is useful	31.546
I was waiting for this notification	15.773
The action demanded by the sender does not require a lot of effort	20.189
At this moment, I was free	37.224

Table 3. User response about why they accept (click) notifications.

Option	Count (%)
Sender is not important	19.565
The content is not important	40.580
The content is not urgent	43.478
The content is not useful	38.406
The action demanded by the sender does require a lot of effort	3.623
At this moment, I was busy	19.565

Table 4. User response about why they dismiss notifications.

UNDERSTANDING THE ACCEPTANCE OF NOTIFICATIONS

In this section we investigate the factors that make the users accept (click) or dismiss a notification. Through the questionnaires, we asked the users the reason for clicking/dismissing a notification (see Table 2). If a notification (linked with the questionnaire) is clicked by the user, we ask them to select all factors that made them decide to click the notification, otherwise, we ask them to select the factors that made them decide to dismiss the notification. We provide a predefined list of seven and six options for clicking and dismissing the notification, in addition, there is a box for open-ended answers in case users do not find an appropriate answer in the provided list.

In Table 3 and Table 4 we calculate the percentage of times each factor was reported as a reason for clicking and dismissing the notifications. Since, users may select more than one option, the count percentage in the table adds up to more than 100%. According to these responses, the users mostly accept notifications when they are free, but also the importance of the sender and the usefulness of the content make them accept a notification. On the other hand, users avoid attending notifications that do not contain important, urgent or useful content. These responses demonstrate that the value of content is used for deciding whether to click or dismiss a notification. Moreover, the user very rarely state that they were busy, thus had to dismiss a notification. This could indicate that the users give precedence to a notification over the primary task, but only if the content is valuable.

Disruptive Notifications are Likely to be Dismissed

We examine the impact of disruption caused by the notifications on their likelihood of being accepted. In order to quantify this, we encoded the response for perceived disruption with the following values: *Strongly disagree*=1, *Somewhat disagree*=2, *Neutral*=3, *Somewhat agree*=4 and *Strongly agree*=5. In order to detect the acceptance of a notification,

Option	Count (%)
Sender is important	25.926
The content is important	33.333
The content is urgent	20.370
The content is useful	35.185
I was waiting for this notification	11.111
The action demanded by the sender does not require a lot of effort	16.667
At this moment, I was free	18.519

Table 5. User response about why they accept disruptive notifications.

we check if it was clicked by the user. In case it was dismissed, we cross-validate the user's response for the question *How did you handle the notification?*. If the user responded that *I decided to dismiss it because it didn't require any further action.*, we mark this notification as accepted. Finally, we use 0 to indicate that the notification is dismissed and 1 for an accepted notification.

We fit a logistic regression model to estimate the effect of perceived disruption on the likelihood of the acceptance of notifications. The model was statistically significant $X^2(1) = 48.3$, $p < 0.001$. The results indicate the likelihood of the acceptance of a notification decreases by 0.581 times (95% confidence interval limits for the slope were [0.497, 0.675]) for a unit increase in the perceived disruption (based on the 5-points Likert scale).

However, the proportion of variance around the fitted model is not high ($R^2 = 0.1434$), which implies that not only the disruption perceived by the user, but also other factors influence the user's decision to accept a notification.

Why disruptive notifications are accepted?

As discussed above, the disruption perceived by the user makes a notification more likely to dismissed. Our dataset shows that 104 out of 474 notifications (with which the questionnaires were linked) were are reported as disruptive. These are the notifications for which the user *somewhat* and *strongly* agreed that they perceived disruption from these notifications.

However, 54 (more than 50%) of these disruptive notifications were accepted (clicked) by the user, regardless of the fact that they caused disruption. To investigate the reason for this, we checked the user's response about the factors that made them click these notifications. Table 5 shows the percentage of times each factor was reported by the users for the disruptive notifications. As users were allowed to select more than one option, the sum of the count percentages in the table adds up to more than 100. "Content is important" and "Content is useful" are the most dominant reason provided by the users for clicking the disruptive notifications. This tells us that even the notifications containing important or useful content can cause disruption. We suspect that these notifications may contain valuable information, but they were not relevant at the moment of delivery. However, we our study remains limited to make such conclusions and provide an understanding about why the users perceive disruption from these useful notifications.

Variable	Estimated Coefficient	Std. Error	t value	p value
Extroversion	0.017481	0.005694	3.070	0.0278*
Agreeableness	-0.012833	0.005387	-2.382	0.0630
Conscientiousness	0.005942	0.004420	1.344	0.2366
Neuroticism	0.008659	0.004042	2.142	0.0851
Openness	-0.003114	0.005369	-0.580	0.5870

N=11
$R^2 = 0.737$
F(5,5)=2.802(p=0.01413)

Table 6. Results of linear regression with the average disruption as a dependent variable, and the five personality traits as independent variables.

DOES PERSONALITY MATTER?

In this section we investigate the role of personality on the reported disruption, seen time and decision time of notifications. We computed the score for the five personality traits (i.e., the so-called Big Five: Extroversion, Agreeableness, Conscientiousness, Neuroticism and Openness) for each of the 11 users who fully completed the exit questionnaire that includes the 50 questions related to personality traits. For this computation, we used the scoring instructions that come with the personality test [18, 5].

Impact on Reported Disruption

To quantify the relationship between the five personality traits and the disruption perceived by the users caused by notifications, we fit a linear regression model with the average disruption as a dependent variable, and the five personality traits as independent variables. Here, the average disruption is computed as a mean of disruption reported by the user through the questionnaires. All responses were encode with the following values *Strongly disagree*=1, *Somewhat disagree*=2, *Neutral*=3, *Somewhat agree*=4, and *Strongly agree*=5. Table 6 shows the parameters of the fitted linear regression model. The results show that the extroversion personality trait significantly affects the average perceived disruption, and that extroverts are more inclined to be disrupted by a notification. The high value for the proportion of variation ($R^2 = 0.737$) in the average disruption depicts that the value of average disruption reported by the users is highly influenced by the personality of users, albeit this could be a consequence of our small sample size.

Impact on Notification's Seen Time and Decision Time

We then investigate whether the personality traits influence the seen and decision time. In order to perform these analyses, we compute:

1. the average seen time of notifications for each user: the average time taken by a user to view a notification. It is computed as the mean of seen time of all notification received by the user.

2. the average decision time of notifications for each user: the average time taken by a user to click/dismiss after viewing a notification. It is computed as the mean of decision time of all notifications received by the user.

Variable	Estimated Coefficient	Std. Error	t value	p value
Extroversion	0.44042	0.14926	2.951	0.0319*
Agreeableness	0.25281	0.14122	1.790	0.1334
Conscientiousness	-0.42050	0.11586	-3.629	0.0151*
Neuroticism	-0.39663	0.10597	-3.743	0.0134*
Openness	-0.05059	0.14074	-0.359	0.7339
N=11				
R^2=0.9007				
F(5,5)=9.073(p=0.01511)				

Table 7. Results of linear regression with the average seen time as a dependent variable, and the five personality traits as independent variables.

Variable	Estimated Coefficient	Std. Error	t value	p value
Extroversion	0.45350	0.13446	3.373	0.0198*
Agreeableness	0.29340	0.12722	2.306	0.0692
Conscientiousness	-0.26102	0.10438	-2.501	0.0544
Neuroticism	-0.36975	0.09547	-3.873	0.0117*
Openness	0.08584	0.12679	0.677	0.5284
N=11				
R^2=0.9035				
F(5,5)=9.366(p=0.01411)				

Table 8. Results of linear regression with the average decision time as a dependent variable, and the five personality traits as independent variables.

We first fit a linear regression model with the average (per user) seen time as a dependent variable, and the five personality traits as independent variables. The parameters of the fitted linear regression model are shown in Table 7. The results demonstrate that the time in which a notification is viewed by the user is significantly influenced by her Extroversion, Conscientiousness and Neuroticism personality traits. We fit another linear regression model with the average decision time as a dependent variable, and the five personality traits as independent variables. The parameters of the fitted linear regression model are shown in Table 8, and demonstrate that the time in which a notification is viewed by the user is significantly influenced by her Extroversion and Neuroticism personality traits.

The need for individual models of human interruptibility has been identified earlier, both in the desktop office setting [21], as well as with mobile smartphone users [30]. The above results show a potential for the interruptibility models to be generalized across groups of users who share the same personality traits.

RELATED WORK

Multitasking is fundamental in workplaces, where task switching happens on a minute scale [19], but also in the private sphere, where an increasing number of personal computing devices mediates the flow of data, be it of entertainment, social or informative nature. Unfortunately, multitasking is seldom seamless, since the limited amount of human attention is sought by a range of competing tasks/themes. The disruptiveness of interruptions was analyzed by Miata and Norman, who were among the first to note and explain its variability with the context in which they take part, and particularly their alignment with respect to the primary task a user is working on [26]. To explain why the interruptions are disruptive, Altmann and Trafton propose *the Memory for Goals model*, which explains how user's intention moves the necessary mental state of the problem between the foreground and the background of a user's attention, and how such state deteriorates when kept in the background [7]. The importance of the problem state held in the memory corresponds to the complexity of the primary task, and consequently, recovering after a complex task is more demanding than if a routine task is interrupted [10].

In this paper, to the best of our knowledge we are the first to investigate the role of the task complexity on interruptibility in the mobile context. Our results confirm the theory of Altmann and Trafton, and we find that interrupting a complex task remains disruptive in the mobile setting. However, the nature of interruptions in our study is fundamentally different than in the above work – our users receive notifications signaling interruptions, and are not "forced" into the interruption per se. This allows us to investigate more subtle phenomena, such as the relationship between the primary task and the time to register a notification. One of our key findings is that *users working on more mentally demanding tasks need less time to notice a notification*. We hypothesize that the "high alert" state in which a user is when working on a complex task [23], also leads to a more agile reaction to a notification.

Multitasking theories were build upon data acquired in highly controlled environments. In reality, however, mobile users engage in unconstrained communication with their friends and family, move about in different surroundings, and get notifications from a range of apps. Unfortunately, datasets of real world mobile notification usage are difficult to obtain. Shirazi et al. collected the data on close to 200 million mobile notifications from 40,000 users [34]. The analysis of the data reveals a high variation of the way a notification is handled depending on the application in is associated with. Personal communication applications, for example, are more preferred and attended faster, than applications associated with utilities and tools. Personal interests, relevance and actionability of the content are additional qualifiers that impact a user's reaction to a notification [16]. In our recent work we uncover that when it comes to communication messages, the relationship between the sender and the receiver plays a significant role in the decision to accept a notification [25]. In this study, we further refine the role of the sender, and find out that the *messages from immediate family members and superior at work are perceived as the least disruptive*. Moreover, the sender is one of the main reasons for accepting a notification, even if the notification interferes with the recipient's current task.

Mobile notifications are arriving in increasing numbers at inappropriate moments threatening to make the whole concept of mobile notifications practically unusable. Automatically inferring the interruptibility of a user and scheduling notifications accordingly is one promising direction for revising the notification mechanism on mobile devices, so that the utility of the notification is preserved. Inferring interruptibility through sensors has been demonstrated in a static office setting by Horvitz et al. [21] and Fogray et al. [17]. In the mobile setting, InterruptMe uses smartphone sensors (GPS, accelerometer), to build an individual model of a user's interruptibility. Contextual cues, such as whether a user is at home or work, the current time of day, whether it is a weekend or not, and the user's level of physical activity, all play role in the user's attentiveness to a notification [30]. Attelia is a system that targets smartphones and smartwatches, and uses built-in accelerometers to recognize moments of task breakpoints [28]. When interrupted at such moments, user's perceive lowered multitasking workload perception. The My Phone and Me dataset goes beyond measuring indicators such as the time to respond and reported disruption level. We observe that not all notifications nor all the receivers are the same, and uncover motivating factors for attending even highly disruptive notifications. The range of significant parameters that determine a user's interruptibility calls for more comprehensive designs and implementations of real-world notification management mechanisms.

LIMITATIONS

Most limitations of the work presented in this paper stem from our decision to collect data in the wild, with the minimum amount of intervention with our users. For example, when it comes to the computation of the seen time of a notification, remotely, we can only detect if a user unlocked the phone and assume that all notifications were seen by her. We cannot detect the precise time when a user starts reading a summary of a message from the notification bar. Moreover, in case a notification arrives when the user is already engaged with the phone, we assume that the user has seen the notification. Further, since our users are not confined to a lab, we are limited to self-reported level of disruption from a notification. In reality, the impact on the primary task need not be high even if the perceived disruption level is high. On the other hand, this self perception might be the most important factor that determines the user's long term sentiment towards notifications. When it comes to our ESM sampling, despite being as light as possible (we ask only up to four ESM questionnaires per day from each user), they increase the number of notifications a user sees during the data collection period. The density of notifications negatively impacts the sentiment towards individual notifications [30], however, we believe that in our case the impact is equally distributed among notifications, and consequently, that the findings about the role of different factors still hold. Finally, while the work is the first to our knowledge to uncover the role of individual psychological traits on mobile interruptibility, it is important to note that we related the traits with the *reported* interruptibility. Therefore, our finding that extrovert people report to be more disturbed by notifications than others, should be interpreted under the fact that general reporting about oneself most likely depends on the person's level of extroversion.

CONCLUSIONS AND FUTURE WORK

In this paper we have presented a study of mobile interruptibility, specifically concentrating on the identification of factors that make an interruption disruptive, and impact the response time to a notification. Through a mixed method of automated smartphone logging and ESM sampling we have obtained a dataset of 10372 notifications and 474 reports on notification perception from 20 users. The analysis of the data reveals factors that impact a user's perception of, and a reaction to, a notification. Thus, the presentation of a notification, its alert type, as well as the type and complexity of a task the user is engaged in, all impact the seen time. Moreover, the relationship with the sender influences the user's decision for accepting a notification or not. The data also reveals how the sentiment (i.e., perceived disruption) towards a notification vary with the type, completion level and complexity of an ongoing task, and the recipient's relationship with the sender. Finally, different people exhibit different reactions, and we observe a substantial role of the individual psychological traits on how a person reacts to a mobile notification.

In future, we plan to advance this work along two fronts. First, we will investigate how the identified factors that impact the disruptiveness of a notification vary in different environments, such as different locations or different times of the day. We hope to develop a holistic hierarchical model of interruptibility that includes subjective and objective sensed attributes. Second, we plan to build upon the orthogonal work that automates interruptibility profiling, and develop a practical tool for detecting suitable moments to send a notification. Our final goal is to automatically infer a user's intention, detect a user's personality from her phone usage, and devise a notification management platform that ensures the notifications are sent out so to align with the user's aims and lifestyle. We believe that considerate notifications will bring us closer to the Weiser's vision of calm technologies that weave themselves into the fabric of everyday life [36].

REFERENCES

1. *Android's Notification Listener Service*. `http://developer.android.com/reference/android/service/notification/NotificationListenerService.html`.

2. *Android's SpeechRecognizer API*. `http://developer.android.com/reference/android/speech/SpeechRecognizer.html`.

3. *Google's Activity Recognition API*. `http://developer.android.com/training/location/activity-recognition.htmll`.

4. *Heads-up Notifications*. `http://developer.android.com/guide/topics/ui/notifiers/notifications.html`.

5. *International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences*. `http://ipip.ori.org`.

6. Piotr D Adamczyk and Brian P Bailey. 2004. If not now, when?: the effects of interruption at different moments within task execution. In *CHI'04*. ACM, Vienna, Austria, 271—278.

7. Erik M Altmann and J Gregory Trafton. 2002. Memory for goals: An activation-based model. *Cognitive science* 26, 1 (2002), 39–83.

8. Brian P Bailey and Joseph A Konstan. 2006. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior* 22, 4 (2006), 685–708.
9. Brian P Bailey, Joseph A Konstan, and John V Carlis. 2000. Measuring the effects of interruptions on task performance in the user interface. In *SMC'00*, Vol. 2. IEEE, Nashville,Tennessee, USA, 757–762.
10. Jelmer P Borst, Niels A Taatgen, and Hedderik van Rijn. 2010. The problem state: a cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, memory, and cognition* 36, 2 (2010), 363.
11. Herbert H Clark. 1996. *Using language*. Vol. 1996. Cambridge University Press Cambridge.
12. Edward Cutrell, Mary Czerwinski, and Eric Horvitz. 2001. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In *Interact'01*. IOS Press, Tokyo, Japan, 263—270.
13. Mary Czerwinski, Edward Cutrell, and Eric Horvitz. 2000a. Instant messaging and interruption: Influence of task type on performance. In *OZCHI'00*. ACM, Sydney, Australia, 361—367.
14. Mary Czerwinski, Edward Cutrell, and Eric Horvitz. 2000b. Instant messaging: Effects of relevance and timing. *People and computers XIV: Proceedings of HCI* 2 (2000), 71–76.
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 *MobileHCI'11*. ACM, Stockholm, Sweden, 181—190.
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 *MobileHCI'10*. ACM, Lisbon, Portugal, 103—112.
17. James Fogarty, Scott E. Hudson, Christopher G. Atkeson, Daniel Avrahami, Jodi Forlizzi, Sara Kiesler, Johnny C. Lee, and Jie Yang. 2005. Predicting Human Interruptibility with Sensors. *ACM Transactions on Computer-Human Interaction* 12, 1 (2005), 119–146.
18. Lewis R Goldberg. 1992. The development of markers for the Big-Five factor structure. *Psychological assessment* 4, 1 (1992), 26.
19. Victor M. Gonzalez and Mark. Gloria. 2004. Constant, Constant, Multi-Tasking Craziness: Managing Multiple Working Spheres. In *CHI'04*. ACM, Vienna, Austria, 113–120.
20. Joyce Ho and Stephen S. Intille. 2005. Using Context-Aware Computing to Reduce the Perceived Burden of Interruptions from Mobile Devices. In *CHI'05*. ACM, Seoul, Korea, 909–918.
21. Eric Horvitz and Johnson Apacible. 2003. Learning and Reasoning about Interruption. In *ICMI'03*. ACM, Vancouver, Canada, 20–27.
22. Shamsi T Iqbal and Eric Horvitz. 2010. Notifications and awareness: a field study of alert usage and preferences. In *CSCW'10*. ACM, Savannah, Georgia, USA, 27–30.
23. Shamsi T Iqbal, Xianjun Sam Zheng, and Brian P Bailey. 2004. Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, Vienna, Austria, 1477–1480.
24. Neal Lathia, Kiran K. Rachuri, Cecilia Mascolo, and George Roussos. 2013. Open Source Smartphone Libraries for Computational Social Science. In *MCSS'13*. ACM, Zurich, Switzerland, 911–920.
25. Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. 2015. Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In *UbiComp'15*. ACM, Osaka, Japan, 813–824.
26. Yoshiro Miyata and Donald A. Norman. 1986. Psychological Issues in Support of Multiple Activities. *User Centered System Design: New Perspectives on Human-Computer Interaction* (1986), 265–284.
27. Christopher A Monk, Deborah A Boehm-Davis, and J Gregory Trafton. 2002. The attentional costs of interrupting task performance at various stages. In *HFES'02*, Vol. 46. SAGE Publications, Baltimore, Maryland, 1824–1828.
28. 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 *Ubicomp'15*. ACM, Osaka, Japan, 475–486.
29. Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing* 16, 1 (2012), 105–114.
30. Veljko Pejovic and Mirco Musolesi. 2014. InterruptMe: designing intelligent prompting mechanisms for pervasive applications. In *UbiComp'14*. ACM, Seattle, WA, USA, 897–908.
31. Martin Pielot, Karen Church, and Rodrigo de Oliveira. 2014a. An in-situ study of mobile phone notifications. In *MobileHCI'14*. ACM, Toronto, ON, Canada, 233–242.
32. Martin Pielot, Rodrigo de Oliveira, Haewoon Kwak, and Nuria Oliver. 2014b. Didn't you see my message?: predicting attentiveness to mobile instant messages. In *CHI'14*. ACM, Toronto, ON, Canada, 3319–3328.

33. Martin Pielot, Tilman Dingler, Jose San Pedro, and Nuria Oliver. 2015. When attention is not scarce-detecting boredom from mobile phone usage. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, Osaka, Japan, 825–836.

34. Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. 2014. Large-scale assessment of mobile notifications. In *CHI'14*. ACM, Toronto, ON, Canada, 3055–3064.

35. 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.

36. Mark Weiser. 1991. The Computer for the 21st Century. *Scientific American* 265, 3 (1991), 94–104.