Context-based Automated Responses of Unavailability in Mobile Messaging

Pranut Jain (corresponding author) (pranutjain@pitt.edu),
Rosta Farzan (rfarzan@pitt.edu) and
Adam J. Lee (adamlee@cs.pitt.edu)
University of Pittsburgh, Pittsburgh, PA, USA

Abstract. People are not always able to respond immediately to incoming messages on their mobile devices, either due to engagement in another task or simply because the moment is inconvenient for them. This delay in responding could affect social relationships, as there are often expectations associated with mobile messaging and people may experience a lingering pressure to attend to their messages. In this work, we investigate an approach for generating automated contextual responses on behalf of message recipients when they are not available to respond. We first identify several types of contextual information that can be obtained from a user's smartphone and explore whether those can be used to explain unavailability. We then assess users' perception of the usefulness of these sensor-based categories and their level of comfort with sharing such information through a Mechanical Turk survey study. Our results show emergent groups with varying preferences with regards to the usefulness and comfort in sharing two types of contextual information: *user state* and *device state*. Further, we also observed a strong influence of message context (i.e., message urgency and social tie strength) in users' perceptions of these auto-generated messages. Our research provides understanding of users' perceptions of sharing context through an autonomous agent that can help design and create effective approaches towards enabling communication awareness.

Keywords: clustering, context-based, messaging, privacy, unavailability, user modeling, utility

1. Introduction

With the advent of internet-based instant messaging applications like WhatsApp[1] and Facebook Messenger[2], mobile messaging has grown tremendously. By 2019, there were over 7 billion IM (Instant Messaging) accounts and this number is projected to rise over 8.9 billion in the next few years[3]. IM has emerged as the preferred method of

[1] WhatsApp, https://www.whatsapp.com/
[2] Facebook Messenger, https://www.messenger.com/
[3] Instant Messaging Statistics Report, https://www.radicati.com/wp/wp-content/uploads/2019/01/Instant_Messaging_Statistics_Report,_2019-2023_Exceutive_Summary.pdf

communication compared to phone call[4] and email[5] [6]. It has gained this popularity due to its informal nature for both social and work related communication (Brown and Barkhuus, 2007). Particularly, in the CSCW community, IM has been described as a successful tool for quick question-answering, as well as for coordination and scheduling due to being less intrusive than phone calls and more immediate than emails while allowing multi-tasking (Nardi et al., 2000; Handel and Herbsleb, 2002; Isaacs et al., 2002). Even though messaging is supposed to be asynchronous, with the advent of mobile devices, studies have shown that most people expect fast responses to their messages (Mai et al., 2015; Pielot et al., 2014). To fulfil these expectations and maintain social relationships, recipients also feel a pressure to respond, even when they are in the middle of ongoing tasks (Avrahami and Hudson, 2004). If only to convey unavailability, message recipients still feel the need to apologize and explain delays in responding (Voida et al., 2002).

These observations point to the importance of awareness in communication. This is also evidenced by recent research in augmenting contextual information in remote communication such as using either heart-rate (Hassib et al., 2017), status messages (Cho et al., 2020), location (Schildt et al., 2016) or a combination of these (Griggio et al., 2019; Buschek et al., 2018) to improve awareness. Dourish and Bellotti, 1992 defines awareness as "understanding of the activities of others, which provides a context for one's own activity". In face-to-face or in-person communication, observing the individual or their surroundings can improve awareness about their activities and availability. However, this effect is diminished to a certain extent in voice calls and even further reduced in IM (Tang, 2007). This lack of awareness makes effective communication difficult, as people often use IM for negotiating or coordinating availability even for communication channels other than IM (e.g., arranging a phone call or in-person meeting) (Nardi et al., 2000; Isaacs et al., 2002; Handel and Herbsleb, 2002). If someone is not reachable over IM, then lack of awareness about unavailability may leave the sender of the message doubtful about how to reach the message recipient. Thus, improving awareness by providing more contextual information in IM not only helps with social

[4] Gallup report, `https://news.gallup.com/poll/179288/new-era-communication-americans.aspx`

[5] GfK MRI Study, https://www.gfk.com/en-us/insights/press-release/smartphone-users-spend-as-much-time-on-entertainment-as-texting-gfk-mri-study/

[6] Flowroute Survey, https://www.flowroute.com/press-type/flowroute-survey-finds-consumers-overwhelmingly-prefer-sms-to-email-and-voice-for-business-interactions/

interactions but also could help the senders' with their own course of action and decision of how to follow up (for example, whether to try another communication channel, or to try a different time, or a different person) (Niemantsverdriet et al., 2019).

Messaging applications like WhatsApp share cues such as the last-seen time, online/offline status, and read receipts to signal availability. However, not only have these cues been shown to be weak indicators of availability, they have also been found to increase social pressures and raise privacy concerns (Pielot et al., 2014; Hoyle et al., 2017; Church and de Oliveira, 2013; Buchenscheit et al., 2014). In other attempts to build availability management systems, Cho et al., 2020 developed a reactive status sharing system for the KakaoTalk app that automatically shared a manually set status for incoming messages to selected close contacts. While this system can share more precise and relevant availability information, it relies on the users to set their status and keep it updated, which they may not do consistently (Begole et al., 2004; Buchenscheit et al., 2014). Rather, Pielot et al., 2014 proposed an automated method of sharing attentiveness level of a user as an indicator of availability. The attentiveness level is predicted using machine learning and utilizing features such as the ringer mode, proximity sensor, and screen activity from a user's smartphone (Pielot et al., 2014). While sharing an attentiveness level was shown to be preferable over existing cues provided by messaging applications, there are drawbacks associated with the attentive prediction as well. A status showing the recipient as *inattentive* can potentially discourage message senders from initiating communication altogether (Dabbish and Kraut, 2004). On the other hand, incorrectly predicting the recipient as *attentive* can have similar negative effect as that of showing that the user is 'online'. Further, just providing an attentiveness status might not be enough as people often need to understand how a technology works and how the attentive/inattentive status was computed to be able to trust it (Abdul et al., 2018).

Another approach presented by Jain et al. involves generating auto-responses when the user is deemed to be *un*available; this can signal inattentiveness but can also utilize model features to explain it (Jain et al., 2019b; Jain et al., 2019a). In that work, authors built personalized attentiveness models to predict unavailability to messaging. They argued that these models can account for individual variations in messaging behavior and can be interpreted to generate contextual explanations. One benefit of the automated responses approach is that these auto-responses can be sent after the sender initiates communication and the recipient is predicted to be unavailable. This way, the sender is not discouraged from starting a conversation by observing a busy flag before even initiating conversation and the recipient will not miss any messages due to that reason. Further, an explanation as to *why* the system predicted the recipient to be unavailable can also boost the sender's trust in the system's prediction (Abdul et al., 2018) and the recipient may not need to explain the delay to the sender later on. However, generating accurate, useful, and trustworthy auto-responses remains an open area of research.

Availability models built by Pielot et al., 2014 and Jain et al. (Jain et al., 2019a; Jain et al., 2019b) used information available directly from a user's smartphone to establish *context*[7] that characterizes the situation of an individual or their device (Dey, 2001). For instance, Pielot et al. used 17 features such as the state of the proximity sensor, ringer mode, and screen setting in their model, while Jain et al. utilized a more comprehensive set of 72 features to represent context at the time of incoming message. These representations of context characterize the environment of the user and their device in rather static form. Dourish, 2004 argues that context is rather an emergent property of the ongoing interaction. That is, (1) not all features are always relevant when accounting for the availability of an individual; i.e. sharing an irrelevant feature as an explanation may not provide any benefit to the message sender, (2) the characteristics of an interaction, such as the purpose of the communication or the relationship between the communicating parties, can influence the shared context utility by affecting how that context is interpreted. Further, these features may contain information that a user might consider sensitive, such as their location. Even if the user is comfortable sharing any type of information, that does not correspond to whether the message sender (with whom the information would be shared) will find that information useful or adequate for explaining unavailability. Thus, it is important to also consider context as a dynamic property of the interaction taking place.

In this work, we analyze people's perception of several types of contextual automated responses from the perspective of both message senders and message recipients. We specifically target one-to-one communication since the expectation of fast responses is more apparent in those type of conversations as opposed to group conversations where a message is usually directed towards multiple conversational participants. A *message recipient* in one-to-one conversation is the person who receives a message from one of their contacts but is unable to respond at that moment, while a *message sender* is the communication

[7] In terms of *collaborative* context, our survey did not provide any explicit information in terms of locality and synchronicity of communication to the respondents. It can be inferred from the examples in the survey that communication is taking place asynchronously in a remote setting.

initiator and the contact who gets the automated response back. We analyze the usefulness of automated responses from the perspective of message senders as well as individuals' comfort level in sharing contextual information from the perspective of the message recipients. Based on our analysis, we then provide design guidelines on how to generate automated responses to manage users' unavailability in responding to mobile messages. Moreover, we provide insight on how different people (both as message senders and recipients) perceive such messages differently and what characteristics contribute to that difference. More specifically, our research addresses the following research questions:

RQ1: What types of automated responses can be generated using contextual information collected from an individual's smartphone?

Auto-responses can be generated in different ways, including simple standard messages, pre-defined 'canned' messages written by users, or messages considering users' current status. In this work, we are particularly interested in messages that are generated based on the context that can be automatically inferred from sensors on people's mobile devices with little or no extra work for the individuals. However, as mentioned earlier, this context can involve a large number of different features. Prior research has identified a large number of features (more than 70) used to represent user context (Pielot et al., 2015; Jain et al., 2019a). Thus, the first step in generating context-based auto-responses involves identifying which types of responses can be put together to create meaningful responses. Therefore, our first research question focuses on identifying the appropriate types of context-based auto-responses.

RQ2: What is the perceived usefulness of different types of automated responses? and How comfortable do people feel with sharing each type of automated response?

For each type of automated response that can be generated, it is critical to understand how message senders (i.e., communication initiators) perceive the usefulness of that response. Additionally, it is important to understand the message recipient's perception about how comfortable they are with the auto response being sent on their behalf. Therefore, our second research question focuses on assessing senders' and recipients' perception of the usefulness and comfort of the context-based auto-responses.

RQ3: How do users differ on their perception of usefulness of and comfort associated with automated responses?

It has been observed that people have varying privacy concerns (Malhotra et al., 2004; Braunstein et al., 2011; Marreiros et al., 2017). Further, people tend to differ on how they utilize technology (Montag et al., 2015; Andone et al., 2016). Personalization is now becoming an integral part in multiple application areas such as web-based systems (Bennett et al., 2012) (Tam and Ho, 2006), learning and education (Bates and Wiest, 2004), banking (Wang et al., 2017) and even availability management (Jain et al., 2019b; Jain et al., 2019a). Thus, it becomes important to consider not only the utility of automated responses but also individual differences which can affect perception of these responses. Identifying characteristics associated with individual preferences can help the autonomous agent adapt to different user groups to address their needs most effectively.

RQ4: What is the role of message urgency and social relationship in the perceived usefulness of and comfort level associated with automated responses?

We hypothesize that communication context in the form of urgency of the message and sender-recipient relationship can play a role in how automated responses are perceived by the users. Multiple previous works have reported the role of relationships in messaging (Mehrotra et al., 2016; Fischer et al., 2010), self-disclosure (Kim et al., 2015; Zhao et al., 2012; Lin and Utz, 2017), location-sharing (Consolvo et al., 2005) and context-sharing (Knittel et al., 2013). Previous works have also pointed out the role of urgency with regards to the reception of communication (Church and de Oliveira, 2013; Teevan and Hehmeyer, 2013). Thus, our fourth research question focuses on understanding how these factors impact users' perception of comfort and usefulness with respect to context-based auto-responses.

To address our research questions, we analyzed a text messaging corpus and conducted a survey through Amazon Mechanical Turk. Our findings indicate varying perceptions about an automated response depending on the context of the information shared through the messages, the relationship with the sender, and the urgency of the message. Our contribution in this work is two-fold: (1) We present the findings of a corpus analysis and how it informed the design and implementation of an online survey about user perception of automated responses; (2) We discuss the implications based on our findings from our survey to design an assistive agent which can support individuals' interpersonal communications through messaging.

2. Background and Related Work

In designing auto-responses for messaging, understanding the issues of expectations, social pressure, and privacy in mobile messaging are of particular importance. In this section, we cover the related literature on these issues. We then discuss related work in developing awareness systems and approaches in evaluating the utility of sharing various categories of contextual information. We highlight the gap in the current research and how our work bridges that gap.

2.1. Availability Management in Communication

Availability Management refers to activities and social processes related to initiating, coordinating, and concluding social interactions (Wiberg and Whittaker, 2005). In face-to-face interactions, individuals rely on social cues, especially in unplanned interactions, to make an informed decision about someone's availability (Tang, 2007). With Computer Mediated Communication, many social cues are no longer present, which can lead to contact initiation at inopportune moments for communication recipients. These initiations have been shown to cause task and social disruptions (Perry et al., 2001; Tang, 2007; Avrahami and Hudson, 2004) while also affecting task performance (Bailey et al., 2001). For example, texting while driving has been attributed as one of the leading causes for automobile related injuries or fatalities (Caird et al., 2014). At the same time, disruptive or ill-timed communications are prone to either being ignored (Chang et al., 2019) or missed (for instance, due to ringer profile set on the phone) (Chang and Tang, 2015; Salovaara et al., 2011). In turn, this can result in negative emotions for message senders. For instance, they might be inclined to negatively adjust their own responsiveness levels towards someone based on how fast their message is responded to (Tyler and Tang, 2003; Pielot et al., 2014).

Internet-based communication applications such as WhatsApp and Facebook compensate the lack of social cues by providing availability indicators such as online status, last-seen, and read receipts (Mai et al., 2015). Mai et al., 2015 through an online survey observed that *intensive* negative emotions are linked to delays in responses specially when senders' are aware that their message has been *seen (or read)* but not responded to. Further, they also observed a higher perceived obligation to respond in message recipients due to the signaling of their message's *seen* status. Similarly, Hoyle et al., 2017 reported based on the result of an online survey the perceptions of message senders when their message is seen but not responded to, with almost 70% respondents reported feeling negative emotions (*upset/angry* or *slighted/ignored*) and 39% **speculated** that they are being ignored or may have been misinterpreted. The authors also reported how recipients were affected by the seen-time, as 68% of survey respondents reported deliberately avoiding viewing a message to pretend not seeing it. On top of this, Pielot et al., 2014 also reported that not only these cues have privacy implications but are not even good indicators of availability emphasizing the need for another method of communicating unavailability.

Social pressure and perceived obligation felt by message recipients can further be inferred by their observed need to apologize and explain delays in responding. For instance, Voida et al., 2002 through interviews, observations, and text analysis identified latent issues attributed to instant messaging. One of the observed behaviors were the need felt by the message recipients to justify delays in responding by providing some situational context possibly as a repair tactic to avoid coming off as rude (from Voida et al., 2002: "talking with Karen...sorry for delay in not talking") (Salovaara et al., 2011). Further, it has been observed that individuals also feel the need to provide context when they need to steer away from a conversation (from Voida et al., 2002: "...I think I'm going to head home right now...can we talk later?") and may even use deceptive or dishonest explanations (Hancock et al., 2009; Reynolds et al., 2013; Salovaara et al., 2011; Voida et al., 2002).

While in synchronous communication methods like phone calls, availability management is implicitly important, the above-mentioned observations not only point to issues surrounding communication but also towards the need to have better awareness mechanisms in messaging as well. In this paper, we build upon these observations regarding the importance of timely responses. Specifically, we explore the possibility of generating context-relevant automatic responses for incoming messages during times when the recipient is found to be unavailable. We show that not only can on-board sensor data used for classifying availability be used to explain unavailability, but also these responses can provide utility to senders without violating the recipient's privacy norms.

2.2. Approaches towards improving awareness in communication

To reduce the impact of interruptions, multiple previous works have focused on accurately detecting inopportune moments for deferring notifications (Mehrotra et al., 2015; Yuan et al., 2017; Pejovic and Musolesi, 2014; Pielot et al., 2017). Similarly, Rosenthal et al., 2011 developed an application which automatically silences the users' phone using predictive models to prevent disruptions. While these approaches try to minimize interruptions and their associated costs, they do not specifically help with enhancing awareness in communication which as discussed in the last section holds important value in social relationships.

In collaborative environments, awareness is deemed important to support cooperation and reduce interruptions to an individual's workflow (Schmidt, 2002). Research in the field of CSCW and computer mediated communication has studied ways to facilitate awareness. Multiple earlier works have explored the use of media spaces that connect remote collaborators with video or audio streams to improve awareness of others' presence and their activities (Harrison, 2009) but have raised concerns related to perceptions surrounding privacy implications (Boyle et al., 2009; Dourish and Bly, 1992). Further, they are confined to limited spaces which are monitored and may become inadequate and harder to track as collaborators move away from their office locations (Bellotti and Bly, 1996).

Lack of easily accessible mechanisms to infer availability may sometimes contribute to unexpected usage of some technologies. Work by Nardi et al., 2000 identified how people use online/offline status of IM applications as a way to infer someone's availability in an office space, which might not always be a good indicator of availability. Tang et al., 2001 extended IM status information for mobile devices to add awareness information such as usage (device idle times) and last used times for calls and IMs. Handel and Wills, 2000 developed a web-based application which compiles information from different sources such as calendars and shared databases (of team information) and presents them to teams to make effective decisions for initiating communication. Wiberg and Whittaker, 2005 developed *The Negotiator* which shared contextual information in the form of availability status (manually set) or preset text messages sharing availability time (e.g., "I will call you back in 0h25m"). The work by Bardram and Hansen, 2010 has explored how context-sharing through mobile devices in a hospital environment improve awareness amongst clinicians. In this case, the shared information included a manual status, activity (from calendar) and location. The information could then be used by clinicians to coordinate shared tasks. More recent work by Buschek et al., 2018 looked at augmenting text messages with additional contextual information such as media playing in the background, local weather, current activities, and distance between the sender and recipient in their prototype application called *ContextChat*. Their work shared context sporadically compared to other works by only augmenting the sent message with contextual information instead of sharing a continuous stream of information. The focus of these works is sharing a static set of contextual information without any modeling or making inferences and allowing the users to make the judgment. This approach might work in some settings such as in a hospital where the clinicians' location and activity provides useful insights in terms their availability to others; however, the approach falls short in less well-defined environments like offices, where location may not provide much insight if employees spend most of their time at their desk. Thus, sharing multiple sources of contextual information not only increases the effort required to interpret them, but may also lead to ambiguous inferences of availability (Begole and Tang, 2007). Further, it is also important to investigate the privacy aspect of sharing multiple contextual information specially if not all the shared information may be relevant in a given situation, which was missing from these works.

Cho et al., 2020 developed a reactive approach for sharing manually set status messages automatically for incoming messages in the KakaoTalk app. Their approach emphasized ease of setting a manual status that will be shared only to incoming messages (similar to an auto-reply) and with fixed contacts. Further, their application allowed message senders to decide whether to alert the message recipient of their message. The issue with this approach is the reliance on users to set their availability status and select the alert type which they might be inconsistent in updating (Begole et al., 2004; Buchenscheit et al., 2014) or may use improperly (e.g., always selecting to alert). This also increases the effort for message senders as they are pinged with not only a status but also another notification asking them to select whether to alert the message recipient or not which could get "annoying" as reported in their user-study. Also, not all status messages may effectively communicate unavailability or be useful to message senders (Knittel et al., 2013) e.g., the use of 'zzz' to denote sleeping may not be interpreted as such by all contacts.

The work by Bardram and Hansen, 2010, Wiberg and Whittaker, 2005 and Cho et al., 2020 included a manual status as part of the shared context. There is an argument that availability management systems need to be automated without requiring manual status input, since typically users are inconsistent in updating their status (Begole et al., 2004; Buchenscheit et al., 2014). Pielot et al., 2014 instead suggested an approach of sharing a predicted user attentiveness level as an indicator of availability. The authors built a model using features such as the *screen activity*, *proximity sensor*, and *ringer mode* which was able to predict a recipient's messaging attentiveness with an accuracy of 71%. They showed that utilizing contextual information which can be captured automatically in a smartphone is a better predictor of availability that WhatsApp's 'last seen' annotation and is also perceived as less privacy invasive. However, there are still some issues that are prevalent with sharing an attentiveness status. For instance, as pointed out in their own user study, respondents felt that an attentiveness status might still create false expectations similar to WhatsApp's and Facebook's existing cues. Respondents also reported not trusting the method since they were not aware of how it was computed (Abdul et al.,

2018) and also questioned the usefulness of knowing the attentiveness level.

Jain et al. proposed a different approach of generating contextual auto-responses to explain unavailability (Jain et al., 2019b; Jain et al., 2019a). By using a more comprehensive feature set of 72 features along with a personalized modeling approach they were able to achieve an accuracy as high as 86% in predicting attentiveness to messaging. The authors hypothesized that an accurate availability model can be interpreted to identify not only inopportune moments for the recipient but also the factors affecting the recipient's availability at that moment. Knowing the reason for unavailability would not only allow for establishing better trust in the recipients' availability state but also could potentially help with better management of expectations. Though, there were some missing discussions related to their work. For instance, in a study by Salovaara et al., 2011, it was pointed out that the unavailability explanations may not need to be truthful but should be 'acceptable'. Thus, not all types of contextual information might be perceived as *useful* when explaining unavailability. In addition to the perceived usefulness of a response, there is also the issue of privacy of the message recipient as there maybe information which the recipient may not be *comfortable* sharing with few, or all of their contacts (Knittel et al., 2013). In this work, we fill these gaps by evaluating the perceptions surrounding both, usefulness, and comfort in sharing of different categories of automated responses considering perceptions of both message senders and recipients. We also explore how the agent can set preferences for the user automatically to reduce their burden.

2.3. Evaluating Utility of Contextual Information

For the design of an interactive awareness system, Niemantsverdriet et al., 2019 proposes the DASS framework and establishes three main themes from prior literature in relation to the development of an awareness system; (1) *What Information is Needed for Awareness?*; (2) *How Can the Awareness Information be Embodied?*; and (3) *How Can the Awareness be Used Effectively in Interaction?*. (1) forms the starting point and further embodies four sub-themes related to the 'type', 'detail', 'inference' and 'privacy' related to awareness information. In particular, Niemantsverdriet et al., 2019 emphasises the importance of understanding the trade-off between usefulness and privacy of the shared information ("when deciding on what information to share, the usefulness of the information needs to be very carefully weighted with the invasion of privacy that is caused by presenting that information to other people"). Exploring this trade-off is a key focus of our work.

Next, we look at works which have evaluated the utility of sharing different types of contextual information to improve awareness in communication.

Khalil and Connelly, 2006 conducted a formative study with 20 participants for a period of 10 days to determine their sharing preferences. Their evaluation included consideration for sharing *four* types of contextual information with *six* different social relations. Their results indicated that participants felt more comfortable disclosing *company* and *in-conversation* than *location* or their *activity* information. Additionally, they considered communication context in terms of social relationship and observed that it was a key factor in the rate of disclosure of different information. Though, this work was more focused towards what the callee is comfortable in disclosing but not the perspective of the caller on whether they find certain contextual information useful. On the other hand, from a caller's perspective, Avrahami et al., 2007 evaluated the effectiveness of different contextual information in allowing callers to make better decisions on when to call or leave a voice message. While they incorporate the importance of urgency, they did not consider social relationship which has been shown as a key factor is self-disclosure (Zhao et al., 2012; Lin and Utz, 2017; Consolvo et al., 2005). Further, the work did not consider callees' privacy considerations as part of their evaluation of different contextual information. Additionally, both these works evaluated a limited set of contextual information and missed some factors which have previously been reported as indicators of availability such as ringer mode, calendar, and device-usage (Chang and Tang, 2015; Jain et al., 2019a; Jain et al., 2019b).

In terms of evaluating a more comprehensive set of contextual categories, De Guzman et al., 2007 conducted a diary study with *13* users over *4* weeks to evaluate *in-situ* perception of contextual categories such as *location*, *time*, *physical availability*, *social availability*, *task status* and *emotional availability.* Similarly, Knittel et al., 2013 through a survey evaluated an even more comprehensive set of contextual information categories such as *location, appointments, activity, phone usage, ringer profile, calling state, app usage, number of surrounding people,* and *mood.* However, some of these categories (e.g., emotional availability, number of surrounding people, etc.) cannot be directly captured from a users' device and might either require prediction or input from the user, which can be distracting or annoying based on how often it needs to be asked or both (Becker et al., 2016). Guzman et al. mentioned that body-worn sensors can be utilized to determine some of these categories which limits the practicality of the approach while also requires making inferences about the users' state which will be ambiguous to a certain degree (Yáñez Cortés, 1975).

Most work in the sharing of contextual information in communication has focused on reducing the disruptions related to phone calls by making callers aware of callees' state before they make a phone call. Thus, it is not clear whether the results observed by these works in relation to usefulness and privacy concerns of sharing different contextual information would also hold when determining the utility of various contextual explanations related to explaining unavailability of message recipients when they are not available to respond. Our work aims to bridge this gap by evaluating the perceptions of different contextual information which can directly be captured through a smartphone in terms of both usefulness and comfort in sharing while also considering communication context.

3. Methods

In this section, we describe the creation of our survey instrument, our survey methodology, and the analysis approaches used to understand peoples' utility and comfort assessments of contextually generated auto-responses in instant messaging platforms.

3.1. Analyzing ways people communicate unavailability

In order to develop an agent to construct contextual auto-responses, the first step is understanding whether and how people typically communicate unavailability. For this purpose, we analyzed an existing text message corpus (O'Day and Calix, 2013). The corpus contains a relatively small number of drug-related criminal messages (labelled) along with a larger number of regular text messages. This corpus is one of the few publicly available messaging corpora with metadata information such as *message time*, *contact id*, and *message type (incoming/outgoing)*. Availability of metadata information makes it easier to identify instances of delayed responses along with their explanations. The corpus contains a total of $4,934$ messages including 289 drug-related messages.

From this corpus, we identified categories of explanations people provide when communicating unavailability. Table 1 lists the identified categories along with examples taken from the corpus. On linking these categories to sensors or features previously used in modeling messaging attentiveness, we established 13 categories of automated responses based on the contextual information they contain. These are listed in Table 2 and represent the categories which we evaluated in our survey. We discuss the analysis of the corpus in more detail in Section 4.1.

3.2. Survey Design

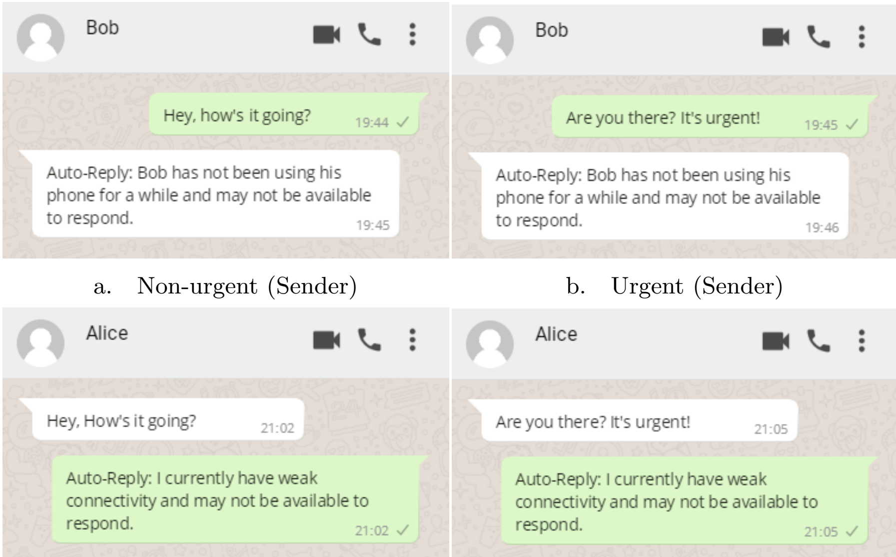
To understand people's perceptions of usefulness and comfort with sharing context-based auto-responses in one-to-one conversations, we designed and conducted a web-based survey [8]. It was distributed using Amazon Mechanical Turk. The study was reviewed and approved by our University's Institutional Review Board. Respondents were paid 3.50 USD for completing the survey. It included two major sections. One of the sections assessed respondents' perception as message senders while the other assessed their perceptions as message recipients about the different types of automated responses listed in Table 2. These were guided by the corpus analysis discussed in section 4.1. Additionally, our survey included questions regarding users' demographic information and privacy concerns.

After the introduction to the survey, participants were presented questions corresponding to message senders' and recipients' perspective in a random order to balance out any potential carry-over effects. Further, the respondents were not informed that they would be asked about the other role and they were not allowed to go back and change their responses after they had completed a section.

3.2.1. *External factors: message urgency and social relationship*

Previous work has shown that a message can be received differently depending on the content of the message and sender-recipient relationship (Fischer et al., 2010; Mehrotra et al., 2016). Two important factors that have been identified by prior work to impact communication and information sharing are urgency of the message and the social relationship between the sender and recipient. Social relationship has been found to have an effect on the willingness to share information (Lin and Utz, 2017; Wiese et al., 2011; Zhao et al., 2012; Consolvo et al., 2005; Knittel et al., 2013; Khalil and Connelly, 2006). Church and de Oliveira, 2013 in their user study pointed out that expectations vary based on the nature of the communication ("If I started a conversation and it's something urgent, then I expect them to respond immediately. If the message isn't important, I personally don't care. I think people respond whenever they find time or whenever they feel like it"). Teevan and Hehmeyer, 2013 also observed that the communication is affected based on if users perceive a communication attempt to be urgent or important. Thus, in the design of our survey we account for the context of communication in the terms of the strength of the social relations

[8] Survey link: `https://people.cs.pitt.edu/~pranut/messaging_study/mstudy_survey.pdf`



a. Non-urgent (Sender) b. Urgent (Sender)

c. Non-urgent (Recipient) d. Urgent (Recipient)

Figure 1. Screen captures distinguished by the urgency. (a) and (b) were shown during the Message Sender's block and (c) and (d) were shown during the Message Recipient's block

and the urgency of the message when evaluating different auto-response types.

3.2.2. *Message Senders Perspective*

Respondents in this section were asked how *useful* they find each category of contextual auto-response on a 3-point scale (Jacoby and Matell, 1971) (3-Useful; 2-Somewhat Useful; and 1-Not Useful). They were presented with four scenarios corresponding to the urgency of their message (i.e., Urgent vs. Not Urgent) and their relationship to the recipient of their message (i.e., Close or frequent contact vs. Distant or infrequent contact). Rather than having fixed relationship groups (such as *friends, families* and *coworkers*) as part of our evaluation, we chose to evaluate the effect of social relations based on closeness and frequency of communication since within social groups, the degree of closeness may vary and closeness has been observed to have a more profound effect than social group on sharing behavior (Wiese et al., 2011). Figure 1a and 1b shows the sample screens that were presented to survey respondents. Here, the top bubble corresponds to their message to the recipient and bottom bubble corresponds to an auto-response. The respondents were asked to rate the usefulness of the auto-response in the four scenarios mentioned above.

After evaluating the *13* auto-response categories, the respondents were asked an open-ended question to provide any additional information from the message recipient's that they would find useful. We also assessed how the granularity of information can influence their judgment of the usefulness of particular messages. For instance, an auto-response message related to a calendar event can include general information about the recipient being in a commitment or include more detailed information of being busy with a *meeting* or a personal event such as *attending a game*. Similarly, auto-responses including location information can include only the general information such as *'not at home'* or *'at work'* or include more detailed information about the exact location.

3.2.3. *Message Recipients Perspective*

In this section, the respondents were asked to assume the role of recipient who receives a message from one of their contacts and were asked to rate how comfortable (3-Comfortable; 2-Neutral; and 1-Not Comfortable) they were with the agent automatically sharing different types of contextual information about their state when they were deemed unavailable. Symmetric to the structure of the message senders block, these questions assessing comfort levels were asked by including communication context (urgency and relation). Figure 1c and Figure 1d show screen samples that the respondents were shown when assessing their comfort levels. The top bubble corresponds the incoming message from a contact and the bottom bubble corresponds to an automated response shared by the agent on the respondent's behalf.

Similar to the sender's block, the questionnaire in this section also included questions related to comfort levels associated with different categories and granularity of shared information within an auto-response category.

3.2.4. *Measuring Privacy Concern*

One assumes that privacy concerns can be an important factor in the design of an auto-response agent, particularly with respect to how comfortable the individuals are with sharing information about their situational context. Therefore, in the third section of the survey, we asked questions relating to the respondent's privacy views to measure their level of privacy concern. Prior work has shown that directly prompting respondents about privacy topics can lead to inflated levels of privacy concern or otherwise biased results (Marreiros et al., 2017; Braunstein et al., 2011). To avoid priming respondents in this manner, we purposefully asked users about their comfort with sharing and utility of auto responses *before* collecting information on general privacy concerns.

To measure a respondent's level of privacy concern, we used the well-established second order IUIPC (Internet Users' Information Privacy Concerns) scale. This scale includes 10 items based on three dimensions - *control* (over information), *awareness* (of privacy practices) and *collection* (of information) (Malhotra et al., 2004). The 10 items are measured on a seven-point scale ranging from *'strongly disagree' (1)* to *'strongly agree' (7)*. Since these set of questions are directly asking about privacy, we expect the responses to be somewhat inflated (Braunstein et al., 2011) but that should not affect our analysis, since we are only interested in measuring *relative* privacy concerns among respondents and to relate that to their responses for usefulness and comfort levels in sharing different information through auto-responses.

3.3. Response Analysis

We utilized a number of statistical analyses in analyzing our survey responses. Here, we describe each analysis approach.

3.3.1. *Factor Analysis*

We evaluated respondents' perception about 13 categories of automated responses in four different contexts, for total of 52 items for both usefulness and comfort. We performed factor analysis to determine if there is a latent structure as to how respondents rated these different categories and if there are some categories or subset of categories that measure the same aspect of perception for an automated response.

Usefulness. Our usefulness response dataset includes responses from 99 respondents for 52 items with each row in the dataset representing a respondent's ratings for each auto-response category. To conduct the factor analysis, we restructured the data into 396 rows, where each row represents the response for specific category of auto-response under unique combination of urgency and social relation values (i.e., *frequent and non-urgent, frequent and urgent, infrequent and non-urgent, infrequent and non-urgent.*

We then performed PCA (Principal Component Analysis) followed by varimax rotation on the transformed data to find components or factors which represent maximum variation in usefulness ratings and to identify any latent structures in how respondents rate different auto-response categories. To find the right number of components, we created the scree plot which compares eigenvalue with different number of components and observed an 'elbow' with two components. The second factor having eigenvalue of 1.090 also satisfies the Kaiser rule of selecting factors > 1.0 eigenvalue. The resulting two components explained 59.996% of the overall variance. Bartlett's test of sphericity was significant ($\chi^2(78) = 2627.052, p < 0.001$). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO value) was 0.941 indicating that the strength of relationships between variables is high.

Comfort. We structured the comfort data in a similar way and conducted a similar factor analysis. Similarly, we observed two components to explain 59.104% of the variation. Bartlett's test of sphericity was significant ($\chi^2(78) = 2652.834, p < 0.001$). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO value) was 0.936 again indicating high strength of relationship between variables.

The factor loadings for both usefulness and comfort are listed in Table 3. The variation captured by component 2 was greater than component 1 in the unrotated components returned by PCA. We interchanged them to better visualize the comparison with the usefulness factor loadings.

3.3.2. *Cluster Analysis*

To assess whether there are groups of respondents who are similar in their responses in terms of their perception of usefulness and comfort, we conducted cluster analysis on the restructured usefulness and comfort datasets with regression scores obtained from factor analysis. We used *k-means* clustering approach with *k-means++* algorithm for selecting initial cluster centers where after randomly selecting the first cluster center from all data points, the subsequent centers are chosen based on probability proportional to the squared distance from existing cluster centers.

To determine the optimal number of clusters k, we used the *elbow* approach by plotting the distortion score associated with different number of clusters. The distortion score computes the sum of squared distances from each point to its assigned center. The test range of k varied from 1 to 6. For both usefulness and comfort datasets, the *elbow* was observed for $k = 3$ which also had the highest silhouette average of 0.461 for usefulness and 0.464 for comfort.

3.3.3. *Regression Analysis*

We performed regression analysis to estimate relationship of respondent attributes (age, gender, etc.) and message context (relation, urgency) with respondent preferences which could be linked to the group they belong to identified from cluster analysis.

Since a respondent rated the usefulness and comfort in sharing of a response category multiple times for each communication context, our dataset includes repeated measures of an auto-response category for each respondent. Therefore, we used *GEE (Generalized Estimating Equations)* which is method used for parameter estimation for correlated data (Liang and Zeger, 1986). In addition to the consideration for dependencies between cases, GEE also does not have distributional assumptions (Pekár and Brabec, 2018).

We used a logistic response model with GEE, with the usefulness or comfort group association as the dependent variable. Each model included respondents' demographics (i.e., age, gender, employment and education) along with self-reported preferred method of communication, frequency of checking for unread messages, IUIPC metrics (i.e., control, awareness and collection) and the message context (i.e., relation and urgency) as the independent variables.

4. Findings

In total, we received 101 responses to our online survey. We removed two responses for failing the *attention check* questions, general low-quality responses (copying question text in open-ended questions) or completing the survey in significantly lower time than the median time of all participants. Our final response set consisted of 99 responses of which 70 respondents reported their gender as *male* (70.71%) and 29 reported as *female* (29.29%). In terms of the age distribution, 46 respondents reported their age between *18-34* (46.46%), 31 between *35-44* (31.31%) and 22 reported greater than *44* (22.22%). Respondents reported their education level as, 16 *high-school or lower* (16.16%), 28 *college or 2-year degree* (28.28%) and 55 *4-year degree or higher* (55.55%). Employment was reported as 84 *employed full-time* (84.85%) and 15 reported *part-time or unemployed* (15.15%).

In terms of preferred method of communication, among our respondents, 58 prefer *messaging* (58.59%), 24 *email* (24.24%) and 17 prefer *calling* (17.17%). We further asked respondents about how frequently they checked their phones for unread messages with 9 reporting *every 5 minutes or less* (9.09%), 51 reporting *couple or more times an hour* (51.52%) and 39 reporting *not more than once an hour* (39.39%).

In terms of *privacy concerns*, the measured concerns among the respondents along all three constructs i.e., *control* ($\mu = 6.077, \sigma = .953$), *awareness* ($\mu = 6.350, \sigma = .840$) and *collection* ($\mu = 5.942, \sigma = 1.007$) were high.

Even though Mechanical Turk worker population in the US has recently been reported to be predominantly male (Difallah et al., 2018), our response-set has larger gender bias towards male population while

Table 1. Categories of explanations identified from the forensics corpus with example and frequencies.

Category	Example	Count
Location	"I am at work"	13
Physical/Motion Activity	"I'm on the way back to campus now"	8
Specific/Other Activity	"I will check later today. I am in a meeting."	10
Sleeping	"Hey sorry took nap"	6
Busy (no context)	"Sorry for not responding, got sidetracked"	7
In conversation	"Still at dinner, in a good conversation. Didn't forget about you."	3
Did not see/notice	"Sorry for getting back to you do late, left my phone on table in other part of house."	7
Weak Connectivity	"I am on the store, getting toilet paper. No reception. What's up?"	2
Low/Dead Battery	"Yeah, my phone died earlier"	3

Table 2. Auto-response categories along with examples.

Category	Example
Busy (no context)	Bob is currently busy and may not be available to respond.
Activity	Bob is currently biking and may not be available to respond.
Connectivity	Bob currently has weak connectivity and may not be available to respond.
Battery Status	Bob's phone is currently low on battery and he may not be available to respond.
Location	Bob is currently at work and may not be available to respond.
Noise Level	Bob is currently in a noisy environment and may not be available to respond.
Charging	Bob's phone is currently charging and he may not be available to respond.
Proximity	Bob's phone is currently covered (in a bag or pocket) and he may not be available to respond.
App Status	Bob is currently playing a game on his phone and may not be available to respond.
Calendar	Bob is currently in a meeting with Joe and may not be available to respond.
Ringer Mode	Bob's phone is currently on silent mode and he may not be available to respond.
Phone Unused	Bob has not been using his phone for a while and may not be available to respond.
Call Status	Bob is currently on a call and may not be available to respond.

other demographics measures are in line with the general Mechanical Turk population[9].

We also checked for order effects in respondent ratings which was insignificant for both i.e., comfort (p =.861) and usefulness (p=.667).

4.1. RQ1: What types of automated responses can be generated using contextual information collected from an individual's smartphone?

On analyzing the messaging corpus described in Section 3.1, we identified 59 explanations that provide situational context for a recipient's

[9] http://crowdsourcing-class.org/readings/downloads/platform/demographics-of-mturk.pdf

unavailability. This includes explaining delays in responding to incoming messages (e.g., “Sorry, just got your text. My phone locked up and i had to do a hard reboot.”); missing an incoming phone call (e.g., “What’s up? Was in church when u called”); or being unable to communicate at the moment (e.g., “I will check later today. I am in a meeting.”). We categorized each message based upon the context provided in the explanation. Table 1 lists the identified categories, with examples of explanations in each category, and the associated count in the corpus. The top recurring context provided in the explanation included location, activity, or physical motion: in 13 cases (22%) the explanations included location-based context, in 10 cases (17%) a specific activity, and in 8 cases (14%) some indication of physical motion. It should be noted that these explanations may not all be accurate or true and may be using deception to politely maintain the social connection (Reynolds et al., 2013; Hancock et al., 2009; Salovaara et al., 2011).

In categorizing the situational context for communicating unavailability, we identified cases like *In conversation*, *Did not see/notice* and *Specific/Other Activity* can have different interpretation depending on more detailed context. For example, the explanation *Did not see/notice*, can be due to different reasons such as the phone is on silent or *DND (Do Not Disturb)* mode or the phone is in a location not being noticed. Similarly, someone can be *In conversation* either face-to-face or on the phone. In generating the messages, however, we posit that such interpretation can be left to the recipient of the auto-response and it is more appropriate for the auto-response to only include the relevant details (Walther and Burgoon, 1992; Earle, 2018). Given the classification, and this assumption, our final auto-response categories are listed in Table 2. These categories can be directly linked to sensor or features used in previous works in modeling messaging attentiveness (Pielot et al., 2014; Jain et al., 2019a). Categories such as *Location*, *Physical Activity*, *Weak Connectivity*, and *Low Battery* can be inferred directly from an individual’s phone sensors whereas more complex categories such as a *Specific/Other Activity*, can either be inferred by the auto-response recipient from an individual’s *Calendar* or the application they are using on their device (*App Status*). Similarly, *Did Not See/Notice* can be inferred from *Ringer Mode*, Last phone use (*Phone Unused*), whether the phone is in pocket/bag (*Proximity*), whether it is *Charging* and what is the surrounding *Noise Level*.

These categories can further be classified based on the information they represent. We define *user-state* categories as those which describe the state or environment of the user while *device-state* categories indicate the characteristics or state of the user’s device. From the categories described in Table 2, *Activity, Location, Noise Level, Calendar, App-Status, Busy (no context)* and *Call-status* would be classified as user-state categories while *Connectivity, Battery Status, Charging, Proximity, Ringer Mode* and *Phone Usage* form the device-state categories. For instance, a user's calendar describes their current schedule and their activity describes their physical state (walking, running etc.). Similarly, while the *Busy* category lacks any additional context, it still describes the user rather than their device. Whereas, battery state and charging categories describe the current power state of the device.

Some categories of explanations have been mentioned in the work by Voida et al., 2002 as well. Quotes from participants in the study included situational context such as *In-conversation* ("talking with Karen...sorry for delay in not talking"), engaged in *another activity* ("...was reading email on my laptop") and *location* (not at home) ("...I'm going to head home right now...can we talk later?). However, Voida et al., 2002 did not categorize the types of contextual indicators included in these explanatory messages, nor explore the possibility of automatically constructing contextual replies. Cho et al., 2020 analyzed the types of manual statuses set by participants to automatically share for incoming messages. They observed six high level categories of statuses set by participants i.e., *Activity*, *Availability*, *Emotional/Physical*, *Location*, *Conversation* and *None*, which are similar to the categories of explanations we identified in our analysis. This further validates our finding into what people already (manually) share when communicating unavailability but also indicates what people might be comfortable sharing automatically on incoming messages. However, their work was limited to close contacts (friends or couples) and did not explore or evaluate the utility of these categories of status messages for message senders, nor how these messages can be generated and shared without manually setting preferences related to each of these categories. We build upon this prior work by first trying to understand (i) the types of context that are useful in explaining unavailability and (ii) the availability of sensor data on the phone to facilitate the creation of auto-responses explaining recipient unavailability.

4.2. RQ2: What is the perceived usefulness and comfort in sharing of different categories of automated responses?

Figure 2 shows the comparison of the usefulness and comfort ratings for all response types. It can be observed that the usefulness ratings of different categories are more spread out than comfort ratings which are

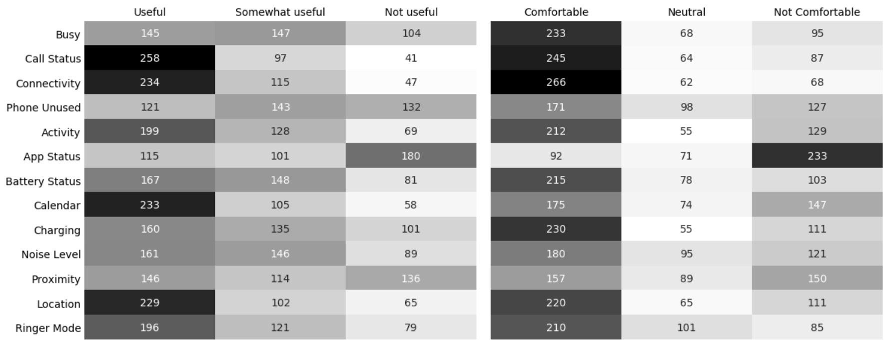


Figure 2. Plot showing overall ratings for different auto-response categories.

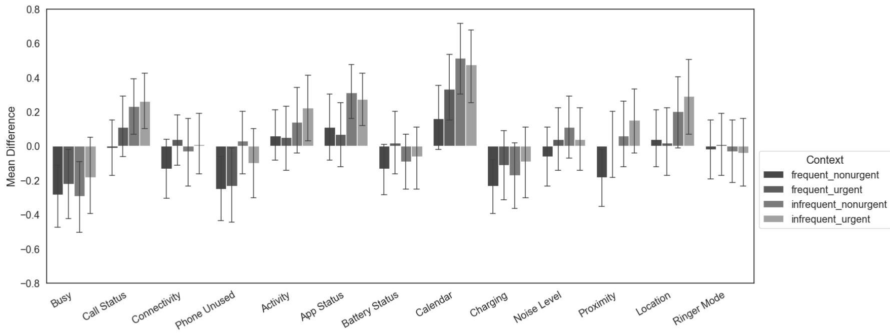


Figure 3. Plot showing the differences between usefulness and comfort ratings for all categories. The error bars represent 95% confidence interval obtained using bootstrapping.

more aligned towards high comfort for most of the categories. This is surprising given the high average privacy concern of our respondents.

Comparing the ratings of different categories, *connectivity* and *call-status* were comparatively perceived as more *useful*, as well rated more *comfortable* in sharing. The utility of recipient's *connectivity* state in particular is an interesting result since previous works on evaluating the value of sharing different contextual information didn't consider the recipients' *connectivity* state since the focus was on reducing disruptions for the callee (Knittel et al., 2013; Khalil and Connelly, 2006). *Read-receipts* in applications such as WhatsApp and Facebook also include a state that represents whether the message has been *delivered* to the recipient (e.g., double white ticks in WhatsApp), though not everyone might be aware or may interpret it as such (Hoyle et al., 2017). This result points to the perceived usefulness of explicitly making the senders' aware of the *connectivity* of the recipient. Further comparing our results to other works, in terms of usefulness, our respondents rated *calendar* and *call-status* higher compared to *phone-usage* whereas in the findings of Knittel et al., 2013 the utility of *phone-usage* was rated higher. In terms of comfort, *location* and *activity* were rated lower compared to *call-status* similar to Khalil and Connelly, 2006. Respondents for the study by Knittel et al., 2013 also rated sharing *App-Status* lower in terms of comfort while *ringer-mode* and *abstract location* had higher disclosure rates pointing to some similarities in terms of comfort of sharing context irrespective of the communication medium. As we will see in Section 4.3, the perception towards different categories varied further based on communication context (social relation and urgency) which was only partially considered in the works by Knittel et al., 2013 and Khalil and Connelly, 2006.

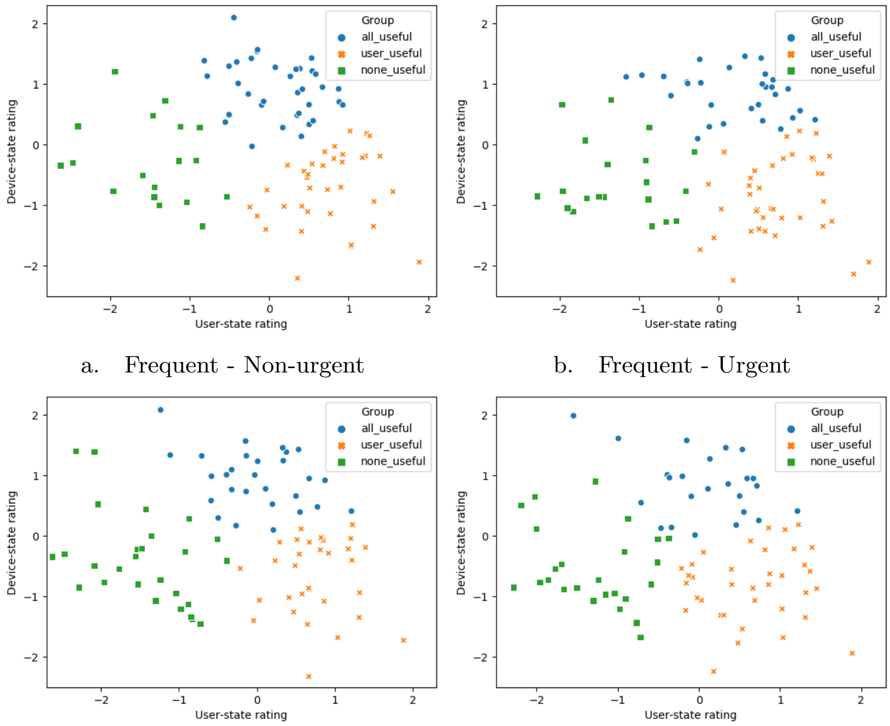
In terms of alignment between usefulness and comfort ratings, Figure 3 visualizes the differences in how respondents rated *usefulness* and *comfort* of different auto-response categories. Directly comparing usefulness and comfort ratings may not give an accurate representation of differences given that Likert scale ratings may not be perceived equidistant from one another by respondents (Sullivan and Artino Jr, 2013). More so, in our survey the middle point for usefulness rating scale was *somewhat useful* which might tend towards positive polarity compared to *neutral* in the comfort rating scale, though the effect this has would be less pronounced than the perceived polarity at the extremes of the Likert scale (Lantz, 2013). Nevertheless, analyzing the mean absolute difference would still give some indication as to where usefulness and comfort ratings differ the most which we observed to be low (i.e., ranging from .50 for *call-status* to .78 for *calendar*). Whereas, the standard deviations were observed to be high (i.e., ranging from .670 for *battery-state* to .771 for *calendar*). On average, categories such as *Busy, Phone-unused, Battery-status* and *Charging* were rated higher on comfort than usefulness, while categories such as *Call-status, Activity, App-status* and *Calendar* were rated higher on usefulness than comfort in sharing indicating existence of varied opinion for some categories with regards to utility and comfort in sharing. Further, the variation between usefulness and comfort ratings for different categories was affected by the communication context (social relation and urgency). For instance, usefulness and comfort associated with sharing *Calendar* is more equally aligned for *frequent* contacts than *infrequent* contacts which relates to a previous finding that relationship category impacts in what way and with whom people share their calendars with (Thayer et al., 2012). These results indicate that our respondents' had varying preferences in terms of perceived utility of different response types and communication context had an effect on their perceptions.

Table 3. Factor Loadings for Usefulness and Comfort ratings.

Usefulness			Comfort		
Category	Component 1	Component 2	Category	Component 1	Component 2
Calendar	.793	.103	Calendar	.847	.140
Call-status	.784	.233	Call-status	.556	.495
Location	.780	.269	Location	.626	.382
Activity	.621	.469	Activity	.711	.326
Busy	.596	.365	Busy	.380	.497
Noise-level	.447	.613	Noise-level	.471	.603
App-Status	.024	.783	App-status	.640	.278
Connectivity	.615	.367	Connectivity	.161	.788
Ringer Mode	.595	.465	Ringer Mode	.468	.538
Battery	.317	.671	Battery	.221	.804
Charging	.423	.664	Charging	.343	.786
Proximity	.328	.763	Proximity	.359	.671
Phone-state	.456	.614	Phone-state	.456	.636

4.2.1. *Variation in preferences based on whether a category represents User-state or Device-state*

The results of factor analysis are presented in Table 3. It can be observed that categories *Calendar, Call-status, Location, Activity* and *Busy* show higher loadings on component 1 than component 2. Whereas, categories *Battery, Charging, Proximity*, and *Phone-state* have higher loadings for component 2 than component 1 for both usefulness and comfort. Most of the categories with higher factor loadings in component 1 represent user related information i.e., user-state categories whereas majority of categories with higher factor loadings for component 2 represent device-state categories as mentioned in Section 4.1. This result suggests how people's perception of potential auto-responses depends on the distinct context of user-related information versus device-related information. Some categories though, such as *App-status* did not correspond to a single component for both usefulness and comfort i.e., the respondents usefulness ratings of app-status were closer to device-state categories than user-state categories whereas for comfort, *App-status* was rated similar to user-state categories. This means that respondents found the usefulness of app-status similar to device-state categories whereas the comfort in sharing was similar to the comfort they felt sharing other user-state categories. Connectivity ratings showed a converse pattern i.e., respondents rated the usefulness of connectivity similar to user-state categories while the comfort in sharing was rated similar to other device-state categories. For the rest



a. Frequent - Non-urgent b. Frequent - Urgent

c. Infrequent - Non-urgent d. Infrequent - Urgent

Figure 4. Scatter plot visualizing user groups based on usefulness ratings for different types of categories identified from Factor Analysis.

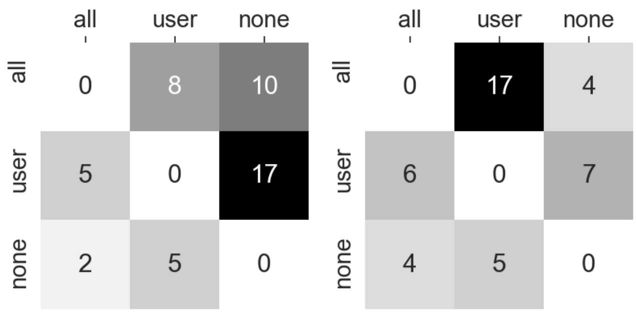
of the paper, we will refer to component 1 as *user-state* categories and component 2 as *device-state* categories.

4.3. RQ3: Emergence of user-groups with varying preferences in relation to the communication context

The standard deviation from the mean for *usefulness* ratings of different categories varied from .675 for *call-status* to .848 for *app-status* and for *comfort* ratings, varied from .771 for *connectivity* to .905 for *activity*. This indicates that, with respect to usefulness of messages and participants' comfort sharing the information, there can be high variation among the respondents. Through cluster analysis described in Section 3.3.2, we identified user-groups with varying preferences for different categories of auto-responses.

Table 4. Number of respondents in each group for different contexts.

Context	all_useful	user_useful	none_useful
Frequent, Non-urgent	44	37	18
Frequent, Urgent	38	41	20
Infrequent, Non-urgent	38	32	29
Infrequent, Urgent	33	37	29



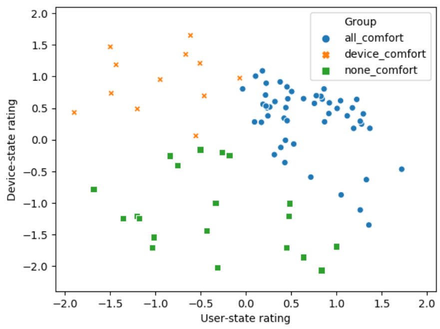
a. Change in relation b. Change in urgency

Figure 5. Changes in respondents' *usefulness* group association (y-axis-*from*, x-axis-*to*) with change in communication context (a. frequent to infrequent and b. non-urgent to urgent). *all* represents *all_useful*, *none* represents *none_useful* and *user* represents *user_useful* groups.

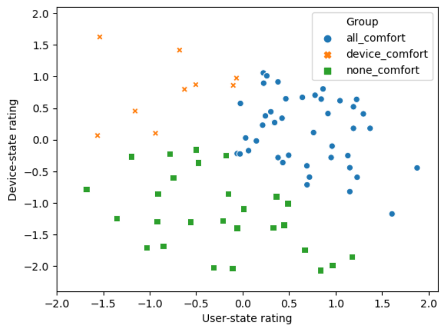
4.3.1. *Usefulness*

Figure 4 shows the plot of the identified clusters in different communication contexts, i.e., *Frequent and non-urgent* (Figure 4a), *Frequent and urgent* (Figure 4b), *Infrequent and non-urgent* (Figure 4c) and *Infrequent and urgent* (Figure 4d). Table 4 lists the number of respondents in each cluster for different contexts.

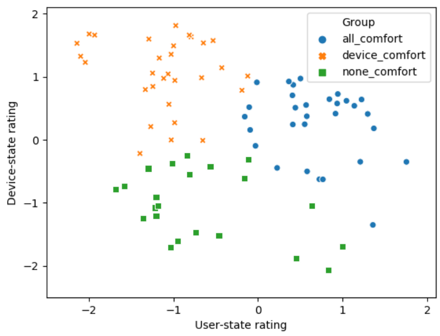
The *x-axis* represents the user-state categories rating and *y-axis* represents the device-state categories rating. Higher values on the *x-axis* represent high rating for user-state categories (activity, location etc.) and higher values for *y-axis* represents higher ratings for device-state categories (phone-status, battery-status etc.). As presented in the plots, one of the emergent groups (depicted in blue dots) has comparatively higher ratings for both user-state and device-state categories which we will refer to as the *'all_useful'* group. Whereas, another group (depicted in green squares) has lower ratings for both user-state and device-state categories which we will refer to as *'none_useful'* group. The third identified cluster (depicted in orange crosses), has higher rating for



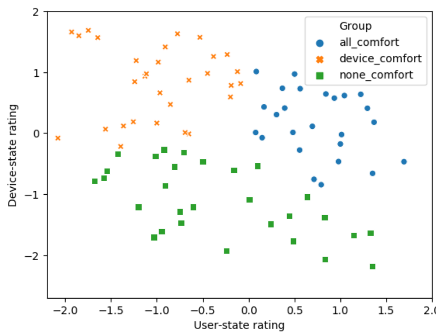
a. Frequent - Non-urgent



b. Frequent - Urgent



c. Infrequent - Non-urgent



d. Infrequent - Urgent

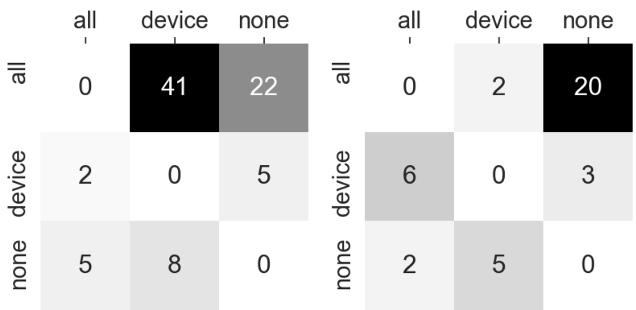
Figure 6. Scatter plot visualizing user groups based on comfort ratings for different types of categories identified from Factor Analysis.

user-state category but lower rating for device-state category which will be referred to as *'user_useful'* group.

Further, we observed that respondents' group association varied based upon the communication context and varying social relation or message urgency resulted in respondents moving from one group to another. Figure 5 shows the change in respondent group association with the change in communication context. For instance, the group association of 27 respondents switched to 'none_useful' group when considering infrequent contacts indicating that for these respondents both types of contextual information (user-state and device-state) was perceived as not useful when trying to communicate with distant contacts. Similarly, 22 respondents switched group association to 'user_useful' group for urgent messages indicating the perceived importance of knowing the user-state in urgent situations.

Table 5. Number of respondents in each group for different contexts.

Context	all_comfort	device_comfort	none_comfort
Frequent, Non-urgent	65	13	21
Frequent, Urgent	58	10	31
Infrequent, Non-urgent	37	32	30
Infrequent, Urgent	30	33	36



a. Change in relation b. Change in urgency

Figure 7. Changes in respondents' *comfort* group association with change in communication context (y-axis-*from*, x-axis-*to*). *all* represents *all_comfort*, *none* represents *none_comfort* and *device* represents *device_comfort* groups.

4.3.2. *Comfort*

Figure 6 shows the plots of the identified clusters in different communication contexts and Table 5 lists the number of respondents in each cluster for different contexts.

Similar to the usefulness plots, the *x-axis* represents the user-state category rating while the *y-axis* represents the device state category rating. The first identified cluster (depicted in blue dots) has comparatively higher comfort ratings for both user-state and device-state sharing which we will refer to as the *'all_comfort'* group, while the second cluster (depicted in green squares) has comparatively lower rating for both user and device state sharing categories which we will refer to as *'none_comfort'* group. The third cluster (depicted in orange crosses) has higher ratings for device-state category while lower ratings for user-state category which we will refer to as *'device_comfort'* group.

For comfort groupings as well, we observed that respondents' group association changed when communication context was varied. Figure 7 shows how the group associations change in different communication contexts. For instance, nearly half of all respondents group association changed from 'all_comfort' to 'device_comfort' group when considering infrequent contacts indicating that these respondents felt comfortable sharing only device-state categories when communicating with distant contacts. Similarly, 23 respondents switched group association to 'none_comfort' group for urgent messages indicating that respondents' were not comfortable sharing both user-state and device-state categories in urgent situations. We elaborate further on this in the discussion section (Section 6).

Overall, there is a bigger shift in group association with context change in comfort ratings compared to usefulness ratings. In general, we observed that respondents' found both user-state and device-state more useful and were more comfortable in sharing those with frequent contacts in non-urgent contexts. Whereas, in urgent contexts it was observed that some respondents perceived user-state information to be more useful. While for infrequent contacts, some respondents rated device-state categories as more comfortable in sharing compared to user-state categories.

These observations indicate that respondents' had varying preferences for different auto-response categories and preferences were affected by the communication context.

4.4. RQ4: Role of user-attributes and communication context on preferences

We observed that communication context affected respondent preferences for different auto-response categories (Section 4.3). The parameter estimates from regression analysis (Section 3.3.3) indicate how significant was the effect of user-attributes and communication context in the perceived usefulness and comfort associated with different auto-response categories.

4.4.1. *Usefulness*

We observed that relation was a significant factor in usefulness group association ($\beta = .373, Exp(\beta) = 1.451, \chi^2(1) = 7.720, p = 0.005$). This suggests that both user and device-state based automated responses are *1.5* times more likely to be found useful when coming from a frequent contact. Employment-status was also marginally significant ($\beta = .913, Exp(\beta) = 2.492, \chi^2(1) = 3.770, p = 0.052$) with full-time employed being *2.5* times more likely to find both user and device-state based automated responses useful.

Neither message urgency nor the interaction effect between social relation and message urgency were significant factors in determining usefulness group association. This implies that message urgency did not significantly affect perception of usefulness for different auto-response categories. Further, other attributes such as gender, age, education and IUIPC were also not significant factors in usefulness group association.

4.4.2. *Comfort*
We observed that gender ($\beta = .846, Exp(\beta) = 2.330, \chi^2(1) = 5.658, p = 0.017$) was a significant factor and male gender was 2.3 times more likely to be comfortable sharing both user and device-state based automated responses. Similar observation was made by Khalil and Connelly, 2006 and Knittel et al., 2013 where men were more likely to share context compared to women. Relation ($\beta = .916, Exp(\beta) = 2.499, \chi^2(1) = 21.350, p < 0.001$) was also a significant factor and respondents were 2.499 times more likely to be comfortable sharing both user and device state auto-responses with frequent contacts. This observation confirms that the importance of social relations in information disclosure (Knittel et al., 2013; Khalil and Connelly, 2006; Zhao et al., 2012; Lin and Utz, 2017; Consolvo et al., 2005) also holds when sharing contextual information through auto-responses. Further, urgency ($\beta = .330, Exp(\beta) = 1.391, \chi^2(1) = 5.030, p = 0.025$) was also a significant factor with respondents being 1.391 times more likely to be comfortable with sharing both user and device-state auto-responses for non-urgent messages compared to urgent messages. We further discuss the implications of message urgency in Section 6.1.

Similar to usefulness, the test for model effects indicated insignificant interaction effect between social relation and message urgency. Other attributes such as age, education, employment, and IUIPC were also insignificant in comfort group association.

5. Predicting Usefulness and Comfort preferences

Automated sharing may raise concerns about unintended or unwilling information disclosure (Wiese et al., 2011). Further, the auto-response sent should also be perceived as 'acceptable' by the sender to effectively communicate unavailability (Salovaara et al., 2011). Thus, it is important to account for user preferences with regards to usefulness and willingness in sharing different auto-response categories. If the user is responsible for setting up their preferences, that would add additional burden of creation and maintenance of policies on the user (Lampinen et al., 2009; Wiese et al., 2011). Rather, for the agent to be accepted, it should be able learn from the users' context and adapt (McDuff and Czerwinski, 2018). Initial preferences for the user can be set based on the group to which they belong in different communication contexts. For instance, given a communication context, if it can be determined

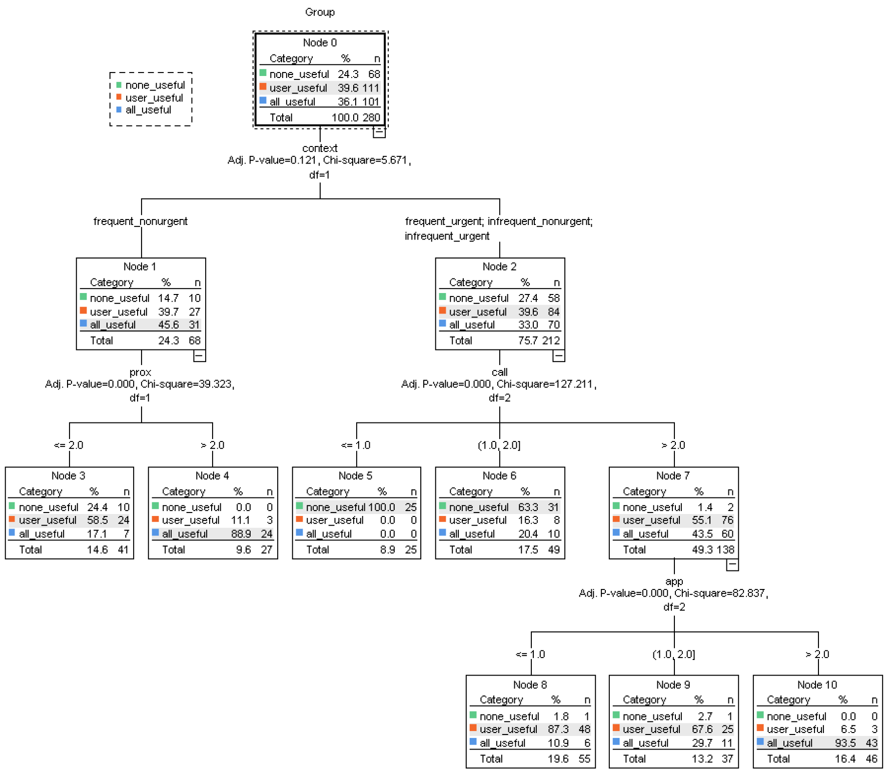


Figure 8. Decision tree visualization for predicting Usefulness group association

that the user belongs to the group 'none_comfort' then it can be implied that they are not comfortable sharing device nor user-state categories in that context.

In Section 4.4, we presented that along with the communication context, the user group association was also affected by demographics such as gender for comfort groups and employment for usefulness groups. Relation strength between contacts can be inferred by looking at the frequency of message exchanges (Wiese et al., 2011) and urgency can either be determined using NLP techniques on messages or utilizing 'important' flag like used in emails (Horvitz and Apacible, 2009). However, demographic information may not always be available, and the collection and/or storage of this information might raise privacy concerns, as this information might be considered sensitive (Awad and Krishnan, 2006). Another way to get initial user preferences is to ask the user to rate all auto-response categories for all communication contexts. This might not only be too cumbersome for the user, but

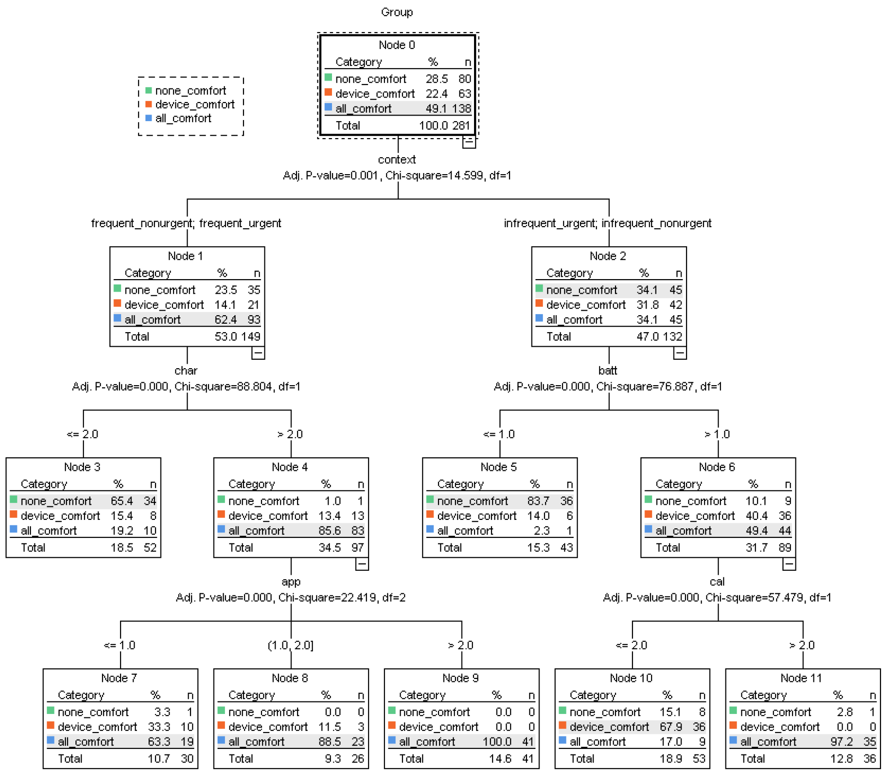


Figure 9. Decision tree visualization for predicting Comfort group association

user preferences might change overtime. Thus, in this section, we evaluate how accurately the usefulness or comfort group association can be predicted and ratings for which categories or subset of categories can best discriminate between different group associations for usefulness and comfort groups.

For this purpose, we built decision trees for the transformed usefulness and comfort ratings response sets with the group associations as the class or ground-truth for each case. One advantage of using decision trees over other classification techniques is easier interpretation and implicit feature selection (Guyon and Elisseeff, 2003). Further, rules can easily be generated by traversing from the root node to the leaf nodes of a decision tree. Since the cases in the usefulness and comfort response sets are not independent, we would need to build 4 models to represent each communication context. Rather than doing that, we split the first or root node of both usefulness and comfort decision trees based on the communication context. We used *CHAID* growing method with 60 minimum cases in parent, 25 in child nodes and *max_depth* of 3 to prevent overfitting. Using a $70 - 30$ training testing split, we got (1) 78.6% training and 74.1% test accuracy for Usefulness group classification with the model asking 3 questions for usefulness rating of proximity, call-status and app-status; (2) 79.7% training accuracy and 75.7% test accuracy for comfort group classification with the model asking 4 questions for comfort ratings of sharing battery-status, charging, calendar and app-status. This indicates that by knowing the message context (urgency and social relation), the agent can make a fairly accurate prediction of an individual's usefulness and comfort group association.

The decision tree models for usefulness and comfort group classification are shown in Figure 8 and 9 respectively. The root node in both trees represents the set of all instances (ratings) used in training phase (280/396). The *usefulness* decision tree splits on three nodes representing *proximity, call-status* and *app-status* while the comfort decision tree splits on four nodes *battery-status, charging, calendar* and *app-status*. Knowing a user's usefulness and comfort in sharing preferences for these subsets of auto-response categories along with the communication context, would allow the agent to predict their initial group associations. As an example, based on the usefulness model, for *frequent contacts in non-urgent* situations, if the message sender rates proximity (device-state category) as *not useful* or *somewhat useful* (≤ 2) they would be classified in the *user_useful* group indicating that they find only user-state categories useful in that communication context. Similarly, for comfort group classification, for *infrequent contacts in non-urgent* situations, if the user rates battery (device-state category) as *neutral* or *comfortable* (> 1) and calendar (user-state category) as *neutral* or *not comfortable* (≤ 2) then they would be classified into *device_comfort* group indicating they are comfortable sharing only device-state categories. As discussed in Section 4.2.1, app-status usefulness was rated similar to device-state categories of component 2 and comfort rating was similar to user-state categories of component 1. This can also be observed with the usefulness group classification tree where one branch represents the rating of app-status and a high rating corresponds to classification to *all_useful* group whereas a low rating corresponds to *user_useful* group i.e., finding categories of user-state (component 1) useful vs finding all categories useful. While the app-value rating does not change the classification of comfort group association, it does affect the confidence associated with the prediction as the number of instances for *all_comfort* group association reduces in the leaf with lower ratings of app-value category.

While this method of classification can initialize a user's preference, further improvements can be made using methods like reinforcement learning where the model can be updated by asking the user to rate the auto-responses as they are sent over time. This can create a more personalized model based on user preferences. Other dimensions of customization in terms of granularity of information in an auto-response and finer or customized contact groups can be considered to further improve the model and understand preferences associated with them. These are beyond the scope of this paper and would be investigated in future work.

6. Discussion

The social information processing theory (SIPT) points to people using any available cues in CMC (Computed Mediated Communication) to make decisions about others and form relations (Earle, 2018; Walther and Burgoon, 1992). Limited or incomplete communication cues may lead to unwarranted speculations in message senders such as *'feeling ignored'* (Hoyle et al., 2017). Providing more relevant context when the recipient may not be available to respond may allow senders' to not only make better inferences about the recipients' state but would also allow for better management of expectations. For instance, when detecting an instance of unavailability for a message recipient who has not been checking their phone, an agent can respond by saying that the recipient *'has not been using their phone for a while'*, this may provide relief to the sender that they are not being ignored, rather the recipient has just not been looking at their phone. When constructing replies such as these, it is important to consider the perceptions surrounding different response types. In particular, message senders should find the information contained in an auto-response useful, and the message recipient should be comfortable sharing this information.

Our work contributes to the growing body of CSCW and HCI research on awareness in remote communication. In particular, our work improves the understanding of automatically acquired context in informing availability. Our analysis identifies important factors for perception and how they can be used to set preferences for individual users. Combined, the results of our work augments the body of knowledge for the design of awareness systems which are cognizant to the communication situation, the type of awareness information, and the preferences of a specific user.

6.1. Design Implications

In this section, we present some implications for the design of an agent-based availability manager based on our findings.

DI1: An agent-based availability manager should be cognizant of user and device state responses. Our findings indicate that the perceptions towards different categories of auto-responses varied based on whether a category represented the *user-state* or the *device-state* contextual information. While we evaluated a limited set of categories based on information that can directly be captured by a user's smartphone, as technology evolves, more information can be made available through additional sensors or in combination with other devices. For instance, multiple respondents noted in the open-ended question asking about other information that they would be comfortable sharing, that they would like the agent to share when they are *'sleeping'* (e.g., "I think I would be comfortable with an auto-response stating that I am sleeping.") which represents user-state and requires making inferences using information from multiple sensors (Chen et al., 2013). The distinction between user and device-state categories would allow adding more response categories—beyond those studied in this paper—without needing to evaluate the utility of every new type of contextual auto-response.

DI2: An agent-based availability manager should account for communication context when determining the type of contextual cue to utilize for communicating unavailability. Another important finding was related to the communication context (i.e., social relationship and message urgency). We observed that relationship has a significant role in both the perceived usefulness as well as comfort in sharing an auto-response category. Our respondents were more likely to share both user-state and device-state based auto-responses with *close or frequent* contacts rather than with *distant or infrequent* contacts (e.g., "I'd be comfortable with just about anything except for people I don't know/talk to often knowing that I might be ghosting them while using my phone like gaming, youtube etc."). Message urgency, too, played a significant role in determining respondents' comfort level associated with both categories of auto-responses, with respondents being more likely to be comfortable sharing both user-state and device-state for non-urgent situations rather than for urgent situations. This observation can be attributed to the fact that people are likely to be more receptive to communication if they perceive it as urgent or important (Avrahami et al., 2007; Teevan and Hehmeyer, 2013; Chang et al., 2019) and would probably like to be able to attend to urgent matters themselves (e.g., "I would be mostly comfortable with anything so long as it isn't urgent. If something were urgent, I would much prefer to be notified about it via some kind of emergency alert rather than an auto response to an urgent message").

DI3: An agent-based availability manager should be aware of individual preferences related to sharing different contextual cues. We also observed individual variations with regards to the perception of auto-responses. As for a given communication context, some respondents were not comfortable sharing any category of auto-responses whereas some were comfortable sharing only device-based context. Similarly, for a given communication context, some respondents found all types of contextual information useful while other found only user-based contextual information useful. The respondents also differed in how much information they would be willing to share with their contacts as some were open to sharing finer details (e.g., "I would be comfortable sharing most any information with close contacts, like who I'm with, where I'm at or what I'm doing. I'd be comfortable with telling my close contacts what time I'll be available again for them to try me again at a more convenient time if I'm doing something I do on a schedule or calendar...") whereas some preferred limiting the amount of details that would be shared (e.g., "The primary concern is that there would be an expectation to respond after I'm done with the activity. So, any activity that is timed and wouldn't take that long to do would be uncomfortable.").

6.2. Practical Considerations

Accuracy of auto-responses: While the focus of this work was to understand perceptions surrounding the utility of different contextual auto-response types, the correctness and accuracy of a response is also important. An accurate availability model can be interpreted to extract relevant information about contextual features that are affecting a user's availability (Jain et al., 2019a; Jain et al., 2019b). These features can then be evaluated to identify which contextual information not only has the most influence in the recipient's unavailability but is also considered comfortable in sharing by the recipient and would be perceived as useful by the sender. The sender's preferences may differ from the recipient's as to what they constitute as useful and need to be either managed/tracked centrally by the agent designer or the agent on sender's phone can send their preferences when communication is initiated.

Burden on recipient: One of the goals of this research is trying to minimize the burden on recipients whether it is due to manually setting an unavailability status or explaining delays in responding to messages. From the point of view of the recipient, they would not have to take any action in case of urgent messages since the urgent signal is directed towards the agent rather than the recipient which then makes the decision on what information to share from the recipient's context (as discussed in Section 5). This is in contrast to prior work by Teevan and Hehmeyer, 2013 and Avrahami et al., 2007 where the urgency context was directed towards the recipient for them to make an informed decision on whether to take the call or not.

Privacy and Mutual Awareness: From a privacy perspective, information should only be shared when communication is initiated, and the recipient is deemed as unavailable. This way, even though more information is being shared about the recipients' context it is limited in terms of accessibility when compared to existing cues in messaging applications and awareness displays proposed in other works (De Guzman et al., 2007; Knittel et al., 2013). Also, since information is only being shared when message senders' initiate communication, the recipient is **aware** of the information that has been shared and with whom (Cho et al., 2020; Niemantsverdriet et al., 2019).

Appropriate use: Finally, as mentioned in Section 4.1, people sometimes use deception in the form of butler lies to signal or explain unavailability (Hancock et al., 2009; Reynolds et al., 2013). People also tend to appropriate technologies to better suit their needs which in terms of mobile messaging might be by turning off 'last seen' (Retore and Almeida, 2019) or by not viewing a message to avoid setting off 'read receipts' (Hoyle et al., 2017). But this tailoring is often based on situational context e.g., contact in question or contacts' setting (personal or professional) and the characteristics of messaging applications (Retore and Almeida, 2019). Whereas manipulating a messaging agent would require understanding of what the agent has already learned about their behavior and what can be done to manipulate the agent behavior to the desired outcome. This might be too complex and could be a limitation in terms of the flexibility of the technology for users to appropriate. Further investigation is needed to assess how people would adapt to the use of a messaging agent for managing unavailability.

6.3. Limitations

Our evaluation of different contextual auto-responses only considered a single response category at a time. It is possible that the combination of multiple categories can have a higher utility in terms of explaining unavailability than individual categories. For instance, *physical activity* information together with *noise levels* may allow the message sender to infer more about the recipients' state. Though, the sheer number of possible combinations would make evaluation overly complex and time-consuming especially for a survey-based study. Further, when asking the respondents to rate a category of contextual auto-reply, while we did mention the possible values that category could represent—e.g., *activity (driving, biking, walking etc.)*—the rating of the respondent may have been biased towards the example which was presented through the sample screen (Figure 1). Further, the middle point for rating categories for usefulness and comfort scales was not the same with the usefulness middle point being *somewhat useful* rather than *neutral*. While this might affect direct comparison between these two dependent variables, it should not affect our analysis of usefulness and comfort individually where the scale for all items was consistent and each rating was considered a distinct class in our analysis.

As mentioned in Section 4, the measured privacy concerns of our respondents were high, which was expected since Mechanical Turk workers on average, have higher privacy concerns than the general US population (Kang et al., 2014). While the measured privacy concerns were not a significant factor in both usefulness and comfort in sharing preferences, the perception towards various categories of contextual information may change with a population with more varied privacy concerns. On the positive side though, even with higher privacy concerns, the respondents were favorable in their perceived comfort associated with sharing different contextual categories indicating the utility of the approach. In terms of generalization, our response-set has a large gender bias towards male population (which is common for Mechanical Turk based studies[10]) and our findings may not be representative of the general US population (Difallah et al., 2018). Finally, our focus with this study was on one-to-one communication. With group messaging the communication dynamics can be different. While it is possible to direct a message sent in a group conversation to a specific individual, group messages are usually intended towards multiple participants and expectation of fast responses or acknowledgements is generally relaxed. The utility of sharing individual context in these situations would require further investigation.

Thus, as future work, we suggest an *'in-the-wild'* study to explore perceptions towards contextual auto-responses generated *'in-situ'* more clearly and identify any additional factors which may represent barriers towards adoption. Another aspect of future work is the perceived value of auto-responses for both message senders and recipients as there

[10] https://www.cloudresearch.com/resources/blog/the-new-new-demographics-on-mechanical-turk-is-there-still-a-gender-gap/

would be additional effort required on the part of the message sender to interpret the received auto-response and make an inference about the recipient's state. If this auto-response is inaccurate or inappropriate, then the recipient would have to put in more effort to remedy the situation. Thus, it is also important to consider what is impact of incorrect decisions made by the agent from the perspective of both message senders and recipients.

7. Conclusion

In this work, we explored an approach towards improving unavailability awareness to message senders through contextual automated responses. We evaluated user perceptions surrounding the usefulness and comfort in sharing various categories of contextual information that can directly be captured using smartphone sensors. Our results showed that the type of context (i.e., user-state or device-state) represented by an auto-response category influenced how it was perceived in terms of usefulness and comfort in sharing. Further, social tie-strength played a role in both usefulness and comfort associated with an auto-response category, while message urgency was relevant in determining the comfort associated with sharing different types of auto-responses. The high variance in ratings for different contextual categories along with presence of distinct groups in terms of perceived utility and comfort ratings of user and device-state categories further point towards variation in individual preferences with regards to perception of auto-responses. We believe that this study will help shape design of better availability management systems which can effectively improve unavailability awareness while considering individual preferences.

Acknowledgements

This work was supported in part by the National Science Foundation under awards CNS–1253204 and CNS–1814866.

References

Abdul, Ashraf; Jo Vermeulen; Danding Wang; Brian Y. Lim; and Mohan Kankanhalli (2018). Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda. In: *CHI '18. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal QC, Canada, 2018. New York, NY, USA, p. 1–18.

Andone, Ionut; Konrad Błaszkiewicz; Mark Eibes; Boris Trendafilov; Christian Montag; and Alexander Markowetz (2016). How Age and Gender Affect Smartphone Usage. In: *UbiComp '16. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, Heidelberg, Germany, 2016. New York, NY, USA, p. 9–12.

Avrahami, Daniel; Darren Gergle; Scott E Hudson; and Sara Kiesler (2007). Improving the match between callers and receivers: A study on the effect of contextual information on cell phone interruptions. *Behaviour & Information Technology*, vol. 26, no. 3, pp. 247–259.

Avrahami, Daniel; and Scott E. Hudson (2004). QnA: Augmenting an Instant Messaging Client to Balance User Responsiveness and Performance. In: *CSCW '04. Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, Chicago, Illinois, USA, 2004. New York, NY, USA, p. 515–518.

Awad, Naveen Farag; and M. S. Krishnan (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. *MIS Quarterly*, vol. 30, no. 1, pp. 13–28.

Bailey, Brian P; Joseph A Konstan; and John V Carlis (2001). The Effects of Interruptions on Task Performance, Annoyance, and Anxiety in the User Interface. In: *INTERACT '01. Human-Computer Interaction: IFIP TC13 International Conference on Human-Computer Interaction, July 9-13*, Tokyo, Japan, 2001. pp. 593–601.

Bardram, Jakob E; and Thomas R Hansen (2010). Context-based workplace awareness. *Computer Supported Cooperative Work (CSCW)*, vol. 19, no. 2, pp. 105–138.

Bates, Eric; and Lynda R. Wiest (2004). Impact of Personalization of Mathematical Word Problems on Student Performance. *The Mathematics Educator*, vol. 14 pp. 17–26.

Becker, Dennis; Vincent Bremer; Burkhardt Funk; Joost Asselbergs; Heleen Riper; and Jeroen Ruwaard (2016). How to Predict Mood? Delving into Features of Smartphone-Based Data. : , *Twenty-second Americas Conference on Information Systems* pp. 1–10.

Begole, James "Bo"; Nicholas E. Matsakis; and John C. Tang (2004). Lilsys: Sensing Unavailability. In: *CSCW '04. Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, Chicago, Illinois, USA, 2004. New York, NY, USA, p. 511–514.

Begole, James "Bo"; and John C. Tang (2007). Incorporating Human and Machine Interpretation of Unavailability and Rhythm Awareness Into the Design of Collaborative Applications. *Human–Computer Interaction*, vol. 22, no. 1-2, pp. 7–45.

Bellotti, Victoria; and Sara Bly (1996). Walking Away from the Desktop Computer: Distributed Collaboration and Mobility in a Product Design Team. In: *CSCW '96. Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work*, Boston, Massachusetts, USA, 1996. New York, NY, USA, p. 209–218.

Bennett, Paul N.; Ryen W. White; Wei Chu; Susan T. Dumais; Peter Bailey; Fedor Borisyuk; and Xiaoyuan Cui (2012). Modeling the Impact of Short- and Long-Term Behavior on Search Personalization. In: *SIGIR '12. Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Portland, Oregon, USA, 2012. New York, NY, USA, p. 185–194.

Boyle, Michael; Carman Neustaedter; and Saul Greenberg (2009). Privacy factors in video-based media spaces. *Media Space 20+ Years of Mediated Life*. Springer, pp. 97–122.

Braunstein, Alex; Laura Granka; and Jessica Staddon (2011). Indirect Content Privacy Surveys: Measuring Privacy without Asking about It. In: *SOUPS '11. Proceedings of the Seventh Symposium on Usable Privacy and Security*, Pittsburgh, Pennsylvania, 2011. New York, NY, USA, pp. 1–14.

Brown, Barry; and Louise Barkhuus (2007). Leisure and CSCW: Introduction to special edition. *Computer Supported Cooperative Work (CSCW)*, vol. 16, no. 1-2, pp. 1–10.

Buchenscheit, Andreas; Bastian Könings; Andreas Neubert; Florian Schaub; Matthias Schneider; and Frank Kargl (2014). Privacy Implications of Presence Sharing in Mobile Messaging Applications. In: *MUM '14. Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia*, Melbourne, Victoria, Australia, 2014. New York, NY, USA, p. 20–29.

Buschek, Daniel; Mariam Hassib; and Florian Alt (2018). Personal mobile messaging in context: Chat augmentations for expressiveness and awareness. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 25, no. 4, pp. 1–33.

Caird, Jeff K.; Kate A. Johnston; Chelsea R. Willness; Mark Asbridge; and Piers Steel (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis & Prevention*, vol. 71 pp. 311 – 318.

Chang, Yung-Ju; Yi-Ju Chung; and Yi-Hao Shih (2019). I Think It's Her: Investigating Smartphone Users' Speculation about Phone Notifications and Its Influence on Attendance. In: *MobileHCI '19. Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, Taipei, Taiwan, 2019. New York, NY, USA.

Chang, Yung-Ju; and John C. Tang (2015). Investigating Mobile Users' Ringer Mode Usage and Attentiveness and Responsiveness to Communication. In: *MobileHCI '15. Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Copenhagen, Denmark, 2015. New York, NY, USA, p. 6–15.

Chen, Zhenyu; Mu Lin; Fanglin Chen; Nicholas D. Lane; Giuseppe Cardone; Rui Wang; Tianxing Li; Yiqiang Chen; Tanzeem Choudhury; and Andrew T. Campbell (2013). Unobtrusive Sleep Monitoring Using Smartphones. In: *PervasiveHealth '13. Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*, Venice, Italy, 2013. Brussels, BEL, p. 145–152.

Cho, Hyunsung; Jinyoung Oh; Juho Kim; and Sung-Ju Lee (2020). I Share, You Care: Private Status Sharing and Sender-Controlled Notifications in Mobile Instant Messaging. *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW1, pp. 1–25.

Church, Karen; and Rodrigo de Oliveira (2013). What's up with Whatsapp? Comparing Mobile Instant Messaging Behaviors with Traditional SMS. In: *MobileHCI '13. Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Munich, Germany, 2013. New York, NY, USA, p. 352–361.

Consolvo, Sunny; Ian E. Smith; Tara Matthews; Anthony LaMarca; Jason Tabert; and Pauline Powledge (2005). Location Disclosure to Social Relations: Why, When, & What People Want to Share. In: *CHI '05. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Portland, Oregon, USA, 2005. New York, NY, USA, p. 81–90.

Dabbish, Laura; and Robert E. Kraut (2004). Controlling Interruptions: Awareness Displays and Social Motivation for Coordination. In: *CSCW '04. Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, Chicago, Illinois, USA, 2004. New York, NY, USA, p. 182–191.

Dey, Anind K (2001). Understanding and using context. *Personal and ubiquitous computing*, vol. 5, no. 1, pp. 4–7.

De Guzman, Edward S.; Moushumi Sharmin; and Brian P. Bailey (2007). Should I Call Now? Understanding What Context is Considered When Deciding Whether to Initiate Remote Communication via Mobile Devices. In: *GI '07. Proceedings of Graphics Interface 2007*, Montreal, Canada, 2007. New York, NY, USA, p. 143–150.

Difallah, Djellel; Elena Filatova; and Panos Ipeirotis (2018). Demographics and Dynamics of Mechanical Turk Workers. In: *WSDM '18. Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, Marina Del Rey, CA, USA, 2018. New York, NY, USA, p. 135–143.

Dourish, Paul (2004). What we talk about when we talk about context. *Personal and ubiquitous computing*, vol. 8, no. 1, pp. 19–30.

Dourish, Paul; and Sara Bly (1992). Portholes: Supporting Awareness in a Distributed Work Group. In: *CHI '92. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Monterey, California, USA, 1992. New York, NY, USA, p. 541–547.

Dourish, Paul; and Victoria Bellotti (1992). Awareness and Coordination in Shared Workspaces. In: *CSCW '92. Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, Toronto, Ontario, Canada, 1992. New York, NY, USA, p. 107–114.

Earle, Kelsey Kathleen (2018), Attributions Online: An Examination of Time Stamps, Read Receipts, and Ellipses in Text-Based Communication. Ph.D. thesis, North Dakota State University.

Fischer, Joel E.; 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. Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, Lisbon, Portugal, 2010. New York, NY, USA, p. 103–112.

Griggio, Carla F.; Midas Nouwens; Joanna McGrenere; and Wendy E. Mackay (2019). Augmenting Couples' Communication with ¡i¿Lifelines¡/i¿: Shared Timelines of Mixed Contextual Information. In: *CHI '19. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, Scotland Uk, 2019. New York, NY, USA, p. 1–13.

Guyon, Isabelle; and André Elisseeff (2003). An introduction to variable and feature selection. *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182.

Hancock, Jeff; Jeremy Birnholtz; Natalya Bazarova; Jamie Guillory; Josh Perlin; and Barrett Amos (2009). Butler Lies: Awareness, Deception and Design. In: *CHI '09. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, USA, 2009. New York, NY, USA, p. 517–526.

Handel, Mark; and Graham Wills (2000). TeamPortal: Providing Team Awareness on the Web. In: *ACM CSCW 2000. Proceedings of the International Workshop on Awareness and the WWW*, Philadelphia, PA, 2000. pp. 3–12.

Handel, Mark; and James D. Herbsleb (2002). What is Chat Doing in the Workplace? In: *CSCW '02. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, New Orleans, Louisiana, USA, 2002. New York, NY, USA, p. 1–10.

Harrison, Steve (2009). *Media Space 20+ Years of Mediated Life*. Springer Publishing Company, Incorporated, 1st edition.
Hassib, Mariam; Daniel Buschek; Paweł W. Wozniak; and Florian Alt (2017), *HeartChat: Heart Rate Augmented Mobile Chat to Support Empathy and Awareness*, p. 2239–2251. New York, NY, USA: Association for Computing Machinery.
Horvitz, Eric J; and Johnson T Apacible (2009), Use of a bulk-email filter within a system for classifying messages for urgency or importance. US Patent 7,565,403.
Hoyle, Roberto; Srijita Das; Apu Kapadia; Adam J. Lee; and Kami Vaniea (2017). Was My Message Read? Privacy and Signaling on Facebook Messenger. In: *CHI '17. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 2017. New York, NY, USA, p. 3838–3842.
Isaacs, Ellen; Alan Walendowski; Steve Whittaker; Diane J. Schiano; and Candace Kamm (2002). The Character, Functions, and Styles of Instant Messaging in the Workplace. In: *CSCW '02. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, New Orleans, Louisiana, USA, 2002. New York, NY, USA, p. 11–20.
Jacoby, Jacob; and Michael S. Matell (1971). Three-Point Likert Scales Are Good Enough. *Journal of Marketing Research*, vol. 8, no. 4, pp. 495–500.
Jain, Pranut; Rosta Farzan; and Adam J. Lee (2019)a. Adaptive Modelling of Attentiveness to Messaging: A Hybrid Approach. In: *UMAP '19. Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, Larnaca, Cyprus, 2019. New York, NY, USA, p. 261–270.
Jain, Pranut; Rosta Farzan; and Adam J. Lee (2019)b. Are You There? Identifying Unavailability in Mobile Messaging. In: *CHI EA '19. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, Scotland Uk, 2019. New York, NY, USA, p. 1–6.
Kang, Ruogu; Stephanie Brown; Laura Dabbish; and Sara Kiesler (2014). Privacy Attitudes of Mechanical Turk Workers and the U.S. Public. In: *SOUPS '14. Proceedings of the Tenth USENIX Conference on Usable Privacy and Security*, Menlo Park, CA, 2014. USA, p. 37–49.
Khalil, Ashraf; and Kay Connelly (2006). Context-Aware Telephony: Privacy Preferences and Sharing Patterns. In: *CSCW '06. Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, Banff, Alberta, Canada, 2006. New York, NY, USA, p. 469–478.
Kim, Bitna; Kyung-Shik Shin; and Sangmi Chai (2015). How People Disclose Themselves Differently According To The Strength Of Relationship In SNS? *Journal of Applied Business Research*, vol. 31, no. 6, pp. 2139.
Knittel, Johannes; Alireza Sahami Shirazi; Niels Henze; and Albrecht Schmidt (2013). Utilizing Contextual Information for Mobile Communication. In: *CHI EA '13. CHI '13 Extended Abstracts on Human Factors in Computing Systems*, Paris, France, 2013. New York, NY, USA, p. 1371–1376.
Lampinen, Airi; Sakari Tamminen; and Antti Oulasvirta (2009). All My People Right Here, Right Now: Management of Group Co-Presence on a Social Networking Site. In: *GROUP '09. Proceedings of the ACM 2009 International Conference on Supporting Group Work*, Sanibel Island, Florida, USA, 2009. New York, NY, USA, p. 281–290.
Lantz, Bjorn (2013). Equidistance of Likert-type scales and validation of inferential methods using experiments and simulations. *The Electronic Journal of Business Research Methods*, vol. 11, no. 1, pp. 16–28.
Liang, Kung-Yee; and Scott L. Zeger (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, vol. 73, no. 1, pp. 13–22.

Lin, Ruoyun; and Sonja Utz (2017). Self-disclosure on SNS: Do disclosure intimacy and narrativity influence interpersonal closeness and social attraction? *Computers in Human Behavior*, vol. 70 pp. 426 – 436.

Mai, Lisa M.; Rainer Freudenthaler; Frank M. Schneider; and Peter Vorderer (2015). "I know you've seen it!" Individual and social factors for users' chatting behavior on Facebook. *Computers in Human Behavior*, vol. 49 pp. 296 – 302.

Malhotra, Naresh K.; Sung S. Kim; and James Agarwal (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, vol. 15, no. 4, pp. 336–355.

Marreiros, Helia; Mirco Tonin; Michael Vlassopoulos; and M.C. Schraefel (2017). "Now that you mention it": A survey experiment on information, inattention and online privacy. *Journal of Economic Behavior & Organization*, vol. 140 pp. 1 – 17.

McDuff, Daniel; and Mary Czerwinski (2018). Designing emotionally sentient agents. *Communications of the ACM*, vol. 61, no. 12, pp. 74–83.

Mehrotra, Abhinav; Mirco Musolesi; Robert Hendley; and Veljko Pejovic (2015). Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In: *UbiComp '15. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Osaka, Japan, 2015. New York, NY, USA, pp. 813–824.

Mehrotra, Abhinav; Veljko Pejovic; Jo Vermeulen; Robert Hendley; and Mirco Musolesi (2016). My Phone and Me: Understanding People's Receptivity to Mobile Notifications. In: *CHI '16. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, San Jose, California, USA, 2016. New York, NY, USA, p. 1021–1032.

Montag, Christian; Konrad Błaszkiewicz; Rayna Sariyska; Bernd Lachmann; Ionut Andone; Boris Trendafilov; Mark Eibes; and Alexander Markowetz (2015). Smartphone usage in the 21st century: who is active on WhatsApp? *BMC research notes*, vol. 8, no. 1, pp. 331.

Nardi, Bonnie A.; Steve Whittaker; and Erin Bradner (2000). Interaction and Outeraction: Instant Messaging in Action. In: *CSCW '00. Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, Philadelphia, Pennsylvania, USA, 2000. New York, NY, USA, p. 79–88.

Niemantsverdriet, Karin; Harm Van Essen; Minna Pakanen; and Berry Eggen (2019). Designing for awareness in interactions with shared systems: the DASS framework. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 26, no. 6, pp. 1–41.

O'Day, Daniel R; and Ricardo A Calix (2013). Text message corpus: applying natural language processing to mobile device forensics. In: *ICMEW 2013. IEEE International Conference on Multimedia and Expo Workshops*, San Jose, CA, 2013. pp. 1–6.

Pejovic, Veljko; and Mirco Musolesi (2014). InterruptMe: Designing Intelligent Prompting Mechanisms for Pervasive Applications. In: *UbiComp '14. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Seattle, Washington, 2014. New York, NY, USA, p. 897–908.

Pekár, Stano; and Marek Brabec (2018). Generalized estimating equations: A pragmatic and flexible approach to the marginal GLM modelling of correlated data in the behavioural sciences. *Ethology*, vol. 124, no. 2, pp. 86–93.

Perry, Mark; Kenton O'hara; Abigail Sellen; Barry Brown; and Richard Harper (2001). Dealing with mobility: understanding access anytime, anywhere. *ACM*

Transactions on Computer-Human Interaction (TOCHI), vol. 8, no. 4, pp. 323–347.

Pielot, Martin; Bruno Cardoso; Kleomenis Katevas; Joan Serrà; Aleksandar Matic; and Nuria Oliver (2017). Beyond interruptibility: Predicting opportune moments to engage mobile phone users. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–25.

Pielot, Martin; Rodrigo de Oliveira; Haewoon Kwak; and Nuria Oliver (2014). Didn't You See My Message? Predicting Attentiveness to Mobile Instant Messages. In: *CHI '14. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Ontario, Canada, 2014. New York, NY, USA, p. 3319–3328.

Pielot, Martin; Tilman Dingler; Jose San Pedro; and Nuria Oliver (2015). When Attention is Not Scarce - Detecting Boredom from Mobile Phone Usage. In: *UbiComp '15. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Osaka, Japan, 2015. New York, NY, USA, p. 825–836.

Retore, Ana Paula; and Leonelo Dell Anhol Almeida (2019). Understanding Appropriation Through End-User Tailoring in Communication Systems: A Case Study on Slack and WhatsApp. In: *SCSM 2019. Social Computing and Social Media. Design, Human Behavior and Analytics*, Cham, 2019. pp. 245–264.

Reynolds, Lindsay; Madeline E. Smith; Jeremy P. Birnholtz; and Jeff T. Hancock (2013). Butler Lies from Both Sides: Actions and Perceptions of Unavailability Management in Texting. In: *CSCW '13. Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, San Antonio, Texas, USA, 2013. New York, NY, USA, p. 769–778.

Rosenthal, Stephanie; Anind K. Dey; and Manuela Veloso (2011). Using Decision-Theoretic Experience Sampling to Build Personalized Mobile Phone Interruption Models. In: *Pervasive'11. Proceedings of the 9th International Conference on Pervasive Computing*, San Francisco, USA, 2011. Berlin, Heidelberg, p. 170–187.

Salovaara, Antti; Antti Lindqvist; Tero Hasu; and Jonna Häkkilä (2011). The Phone Rings but the User Doesn't Answer: Unavailability in Mobile Communication. In: *MobileHCI '11. Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, Stockholm, Sweden, 2011. New York, NY, USA, p. 503–512.

Schildt, Emily; Martin Leinfors; and Louise Barkhuus (2016). Communication, Coordination and Awareness around Continuous Location Sharing. In: *GROUP '16. Proceedings of the 19th International Conference on Supporting Group Work*, Sanibel Island, Florida, USA, 2016. New York, NY, USA, p. 257–265.

Schmidt, Kjeld (2002). The problem with awareness': Introductory remarks on-awareness in CSCW'. *Computer Supported Cooperative Work (CSCW)*, vol. 11, no. 3-4, pp. 285–298.

Sullivan, Gail M; and Anthony R Artino Jr (2013). Analyzing and interpreting data from Likert-type scales. *Journal of graduate medical education*, vol. 5, no. 4, pp. 541–542.

Tam, Kar Yan; and Shuk Ying Ho (2006). Understanding the Impact of Web Personalization on User Information Processing and Decision Outcomes. *MIS Quarterly*, vol. 30, no. 4, pp. 865–890.

Tang, John C (2007). Approaching and leave-taking: Negotiating contact in computer-mediated communication. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 14, no. 1, pp. 5–es.

Tang, John C.; Nicole Yankelovich; James Begole; Max Van Kleek; Francis Li; and Janak Bhalodia (2001). ConNexus to Awarenex: Extending Awareness to Mobile

Users. In: *CHI '01. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Seattle, Washington, USA, 2001. New York, NY, USA, p. 221–228.

Teevan, Jaime; and Alexander Hehmeyer (2013). Understanding How the Projection of Availability State Impacts the Reception Incoming Communication. In: *CSCW '13. Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, San Antonio, Texas, USA, 2013. New York, NY, USA, p. 753–758.

Thayer, Alexander; Matthew J. Bietz; Katie Derthick; and Charlotte P. Lee (2012). I Love You, Let's Share Calendars: Calendar Sharing as Relationship Work. In: *CSCW '12. Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, Seattle, Washington, USA, 2012. New York, NY, USA, p. 749–758.

Tyler, Joshua R.; and John C. Tang (2003). When Can I Expect an Email Response? A Study of Rhythms in Email Usage. In: *ECSCW 2003. Proceedings of the 2003 European Conference on Computer-Supported Cooperative Work*, Dordrecht, 2003. pp. 239–258.

Voida, Amy; Wendy C. Newstetter; and Elizabeth D. Mynatt (2002). When Conventions Collide: The Tensions of Instant Messaging Attributed. In: *CHI '02. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Minneapolis, Minnesota, USA, 2002. New York, NY, USA, p. 187–194.

Walther, Joseph B.; and Judee K. Burgoon (1992). Relational Communication in Computer-Mediated Interaction. *Human Communication Research*, vol. 19, no. 1, pp. 50–88.

Wang, May; Stella Cho; and Trey Denton (2017). The impact of personalization and compatibility with past experience on e-banking usage. *International Journal of Bank Marketing*, vol. 35, no. 1, pp. 45–55.

Wiberg, Mikael; and Steve Whittaker (2005). Managing availability: Supporting lightweight negotiations to handle interruptions. *ACM transactions on computer-human interaction (TOCHI)*, vol. 12, no. 4, pp. 356–387.

Wiese, Jason; Patrick Gage Kelley; Lorrie Faith Cranor; Laura Dabbish; Jason I. Hong; and John Zimmerman (2011). Are You Close with Me? Are You Nearby? Investigating Social Groups, Closeness, and Willingness to Share. In: *UbiComp '11. Proceedings of the 13th International Conference on Ubiquitous Computing*, Beijing, China, 2011. New York, NY, USA, p. 197–206.

Yuan, Fengpeng; Xianyi Gao; and Janne Lindqvist (2017). How Busy Are You?: Predicting the Interruptibility Intensity of Mobile Users. In: *CHI '17. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 2017. New York, NY, USA, pp. 5346–5360.

Yáñez Cortés, R (1975). The problem of interpretation in psychology. *Acta psiquiatrica y psicologica de America latina*, vol. 21, no. 2, pp. 84—89.

Zhao, Chen; Pamela Hinds; and Ge Gao (2012). How and to Whom People Share: The Role of Culture in Self-Disclosure in Online Communities. In: *CSCW '12. Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, Seattle, Washington, USA, 2012. New York, NY, USA, p. 67–76.