sciendo

Proceedings on Privacy Enhancing Technologies ; 2020 (2):416–435

Daricia Wilkinson*, Paritosh Bahirat, Moses Namara, Jing Lyu, Arwa Alsubhi, Jessica Qiu, Pamela Wisniewski, and Bart P. Knijnenburg

Privacy at a Glance: The User-Centric Design of Glanceable Data Exposure Visualizations

Abstract: Smartphone users are often unaware of mobile applications' ("apps") third-party data collection and sharing practices, which put them at higher risk of privacy breaches. One way to raise awareness of these practices is by providing unobtrusive but pervasive visualizations that can be presented in a glanceable manner. In this paper, we applied Wogalter et al.'s Communication-Human Information Processing model (C-HIP) to design and prototype eight different visualizations that depict smartphone apps' data sharing activities. We varied the granularity and type (i.e., data-centric or app-centric) of information shown to users and used the screensaver/lock screen as a design probe. Through interview-based design probes with Android users (n=15), we investigated the aspects of the data exposure visualizations that influenced users' comprehension and privacy awareness. Our results shed light on how users' perceptions of privacy boundaries influence their preference regarding the information structure of these visualizations, and the tensions that exist in these visualizations between glanceability and granularity. We discuss how a pervasive, soft paternalistic approach to privacy-related visualization may raise awareness by enhancing the transparency of information flow, thereby, unobtrusively increasing users' understanding of data sharing practices of mobile apps. We also discuss implications for privacy research and glanceable security.

Keywords: privacy, user, design, qualitative, mobile

DOI 10.2478/popets-2020-0034
Received 2019-08-31; revised 2019-12-15; accepted 2019-12-16.

***Corresponding Author: Daricia Wilkinson:** Clemson University, E-mail: dariciw@clemson.edu
Paritosh Bahirat: Clemson University, E-mail: pbahira@g.clemson.edu
Moses Namara: Clemson University, E-mail: mosesn@g.clemson.edu
Jing Lyu: Clemson University, E-mail: jlu5@g.clemson.edu
Arwa Alsubhi: Clemson University, E-mail: aalsubh@g.clemson.edu
Jessica Qiu: Clemson University, E-mail: jessicq@g.clemson.edu
Pamela Wisniewski: University of Central Florida, E-mail: Pamela.Wisniewski@ucf.edu
Bart P. Knijnenburg: Clemson University, E-mail: bartk@clemson.edu

1 Introduction

In the United States, 77% of the population owns a smartphone [14]. Mobile phone users heavily rely on mo-bile applications (or "apps") to perform everyday tasks, ranging from cultivating social relationships [12], getting to places they want to go [52], to managing their personal productivity [6]. To take advantage of these conveniences, users often have to trust these apps enough to give them access to their personal information, such as their personal contacts, physical location, and calendar events [64, 70]. While there are definite benefits to sharing personal information with mobile smartphone apps [41], there are also concerns as to whether apps collect more information about users than needed [77] and whether this information harvesting is ethical and transparent to users [22].

To add to this problem, developers often leverage third-party libraries to improve user engagement, analytics, and advertising [7]. These libraries often serve as aggregator services that centralize personal information gathered across multiple apps [16, 20, 59], which is then sold for profit [2]. This data could potentially be used to make inferences about users' behavior, socio-economic status, and even their political leanings [11]. While this has direct consequences for smartphone users' online privacy, users have a difficult time knowing who has this data and what is being done with it [2], since there are no visual cues that indicate if or when this data is being shared.

In response to this problem, researchers have developed means to identify applications that leak Personally Identifiable Information (PII) and methods to help reduce risk exposure (e.g., [4, 29, 57, 59, 66]). This effort contributes significantly to an academic understanding of privacy leaks and violations in the mobile ecosystem; yet, most of these existing solutions tend to focus on the accuracy of detection, without considering

whether users actually understand what is being presented. Moreover, even a comprehensively usable solution to this problem would have to rely on users' motivation to access the information and take subsequent action regularly. Given the well publicized gap between users' stated privacy concerns and their subsequent actions [18, 47], one can predict that all but the most privacy-concerned users are likely to avoid the required effort to do this in a way that is sustainable over time.

Arguably, users could become more aware of the data sharing practices of their smartphone applications if privacy information were to be provided in a pervasive but unobtrusive manner, for instance, by adding this information to their screensaver or lock screen. The goal of our work is to develop pervasive but unobtrusive visualizations that enhance users' understanding of the real-time data-sharing practices of apps installed on their mobile devices. We conduct a user study evaluating different design prototypes of these visualizations to investigate the following research questions:

RQ1: Are users able to understand the purpose of the privacy visualizations and the data flows depicted by the visualizations at-first-glance? If so, what contributes to or hinders their understanding?

RQ2: Does a more prolonged inspection of the privacy visualizations increase users' understanding of data sharing practices? If so, what contributes to or hinders their understanding?

RQ3: Do users prefer more "application-centric" or "data-centric" designs? If so, why?

To answer these research questions, we administered a design probe and interviewed 15 participants who each evaluated two of our eight different design prototypes that varied by information structure (i.e., "app-centric" versus "data-centric") and the granularity of information. Our main contribution is a series of glanceable data exposure visualizations that communicate data exposure. By manipulating the structure and granularity of the information, we gained insights into users' preference regarding these aspects, as well as the effect these aspects have on participants' understanding of the visualizations and the information they contain. We found that our participants' preference of information structure depended on their perceptions of privacy boundaries as characterized by Petronio's theory of Communication Privacy Management (CPM) [53]; participants who considered apps to be appropriate co-owners of their personal information preferred the app-centric designs, whereas those who were more focused on the information being shared (regardless of the app) preferred the data-centric designs. These findings imply that there is an opportunity for the design of attention-sustaining visualizations that are aligned with users' specific perceptions of privacy boundaries.

This study contributes to privacy literature by taking a novel approach to visualizing data flows in a uniquely pervasive, yet unobtrusive way, while exploring the effect of the content of such visualizations on users' understanding of the privacy risks posed by their mobile apps. Focusing on U.S. Android users, this study provides insight into how developers and researchers can better inform mobile users about their privacy in an unobtrusive manner.

In the following sections, we begin with our background motivation and study methodology, followed by a description of our design rationale and details of our in-person user study. We conclude by discussing the potential implications of our findings and plans for future work.

2 Motivation and Background

In this section, we discuss the existing literature from various perspectives. We begin by discussing general concerns associated with privacy in Android smartphones and different approaches researchers have taken to understand user awareness, mental models, and attitudes towards managing privacy on Android smartphones. Furthermore, we discuss research on detecting data leaks and on alternative means of presenting data leak information to users. Finally, we discuss literature about glanceable designs from whence we draw our motivation of glanceable privacy.

2.1 Informed Consent, Awareness and Risk Communication

Online privacy can be viewed as the right to limit access to one's personal information [65]. However, in an era of hyper-connectivity and the massive collection of very detailed information (e.g., their location, their Internet searches they make, their heart rate), it becomes difficult for users to understand and manage the data collection and sharing practices of the apps on their phones [21]. The main motivation of this work is to examine the beliefs, expectations, and concerns of users to understand the relationship between them and the data economies they interact with.

Smartphone owners have a great responsibility to be attentive to risk communications. Apps usually provide lists of requested permissions—a summary of what type of data the app will be collecting. However, research suggests that these permissions are ineffective because users do not pay attention to them [9]. Particularly, in one study only 17% of users reported having paid attention to permission lists when downloading an app while 42% of users were not aware of permission lists in general [21]. Studies have shown that smartphone users simply click through notifications or warnings [10, 12, 15–17]—a clear sign of habituation—which negatively affects their awareness of possible risks. Research also indicates that even when Android users read through the permissions presented by the app developers, they may not necessarily understand what these permissions really mean [33].

In sum, users are often inattentive to risk communication, and it may be difficult for them to understand and make informed decisions about potential privacy risks that are communicated to them by the app developer [3]. Worse yet, research has demonstrated that there exist many risks to users' privacy that are not communicated by the developer. For example, Zang et al. [78] found that 73% of 110 popular free Android apps shared personal information such as email addresses with third parties, and a significant number of those apps are not required to notify users when they share such data with third parties.

Many apps mention their data collection policies in their privacy policies, but studies have shown that these policies are often dense, not easily understood, and ignored by users [32, 73]. Thus, researchers have explored privacy awareness tools aimed at educating users about data collection. Previous studies have explored using privacy nudges [72], privacy facts displayed as nutrition labels [32] and visual cues that serve as a warning of suspicious activity [55]. Researchers have also examined changes to the permission interface on Android's Google Play Store by using simplified language [34] or more detailed explanations about information disclosure [37].

Although there have been considerable strides in working towards usable privacy solutions in this domain, there are several significant limitations. Typically, users are not permitted granular control to manage data access, use, or sharing. Rather, their control is limited to broad source types such as "contacts" or "photo album". Moreover, when such broad strokes permissions are granted, users are not notified about *what* information is actually transmitted to outside servers, *when* this happens, and *which app* transmits this information. More recent work has explored different techniques for providing more comprehensive, granular, real-time detection of such "privacy leaks" [4, 20, 29, 57, 58, 66]. A full coverage of the technical details is outside the scope of this paper—an overview of the techniques used in these detection apps can be found in [47].

2.2 Data Leaks and Privacy Visualizations

User research pertaining to privacy aspects of Android OS and associated applications revolves around understanding how users perceive permission requests. For example, Fu et al. [23] conducted a user study to explain the gap between user understanding of location-related permissions and the actual data collection by an app. They found that users varied greatly in their understanding of 'Approximate' location as opposed to 'Precise' location.

Other research on Android OS privacy employs Machine Learning methods to support users' decision-making regarding privacy permissions. For example, Liu et al. conducted an analysis of user decisions to develop profiles that can assist in simplifying permission requests [45]. Liu et al. also developed a personal privacy assistant to assist users with decisions related to Android permissions [44]. Oglaza et al. propose a recommender-based privacy management system which learns from users' privacy preferences to propose authorization rules in order to manage permissions [50].

Researchers have also proposed architectures that give Android users more flexibility in managing their privacy permissions and that help them in understanding the frequency and timing of data access by various applications on their phones [61, 62]. Their work has focused on the detection of privacy leaks and on improving end-user awareness of privacy leaks, but still lacks significantly when it comes to presenting this crucial information to the end user in an understandable way.

In this regard, researchers acknowledge that there are several ways of presenting the information and that the optimal presentation should be investigated [30, 73]. Indeed, research aimed at understanding users' privacy behavior to develop usable privacy-setting interfaces that present information in an appropriate manner has been proven helpful in social networking (specifically Facebook). Notable work in this area includes Lipford et al. who developed an "audience view" that tells users how their profile will appear to other users based on the settings which have been applied [43].

2.3 Glanceable Design to Enhance User Engagement

Current privacy control and awareness solutions, primarily rely on user motivation to first open an app, review the information being accessed and subsequently take action if they are uncomfortable with the finding [39, 57, 66]. Unfortunately, research shows that despite expressing interest in full control over their private information users are less motivated to actually act on their concerns [18, 49]. To motivate users, some "nudge" or "trigger" should be employed to urge them to investigate whenever there is a significant number of privacy leaks.

On smartphones, the most common trigger mechanism is a notification. Prior work on smartphone usage finds that usage is heavily dominated by "checking" habits that regularly act as "gateways to other applications and opening portals to dynamic content" [51, 69]. However, providing notifications every time a piece of data is shared can be perceived as intrusive or annoying by the users and eventually leads to inattention or habituation [54, 74].

Similarly, research on activity trackers, finds that glanceable behavioral feedback interfaces can enhance user engagement and have a positive impact on user behavior [25]. Specifically, Gouveia et al. [24], found that the use of activity trackers is dominated by brief 5-second glances i.e., users call on an application to check up on their current activity levels without any further interaction. Subsequently, they leveraged this finding to explore designs that would increase the frequency of glances to positively impact user physical activity, and promote moments of exploration and learning with activity trackers. Likewise, Klasnja et al. [35] also showed that presence of glanceable displays enables users to keep themselves engaged in physical activities. In our current work, we leverage these findings such as "brief 5-second glances" to examine visualization designs of "glanceable" displays that disclose ongoing data sharing practices to motivate users to take actions that would limit app access to their private information as they deem necessary. We posit that these kind of visualizations would leverage glances to positively motivate users to safeguard their privacy.

In the subsequent sections, we discuss the user-centric design and evaluation of a potential solution to users' lack of awareness about data exposure caused by Android apps in the form of pervasive and unobtrusive data exposure visualizations. These visualizations leverage glanceable design to motivate smartphone users to manage their privacy and to empower them to make informed privacy decisions. We also discuss how our visualizations can be leveraged in mobile devices by integrating them into users' screensavers or lock screens.

3 Communication-Human Information Processing as a Model for Privacy Design

The purpose of presenting warnings or notices to users about their privacy is to influence their behavior: if users are warned about apps that engage in unwanted information sharing, they may restrict their permissions or even remove those apps altogether. However, the information presented in warnings or notices has to go through several stages before it can influence a user's behavior [38]. Prior work in cognitive psychology and the science of warnings provides useful information on how people process and react to warnings [63]. Specifically, Wogalter et al.'s Communication-Human Information Processing Model (C-HIP) [76] for risk communications gives insight into how people process information in terms of noticing it, understanding what is presented, determining its significance and if action is needed (see Figure 1). Our designs are influenced by each stage of this framework as described below.

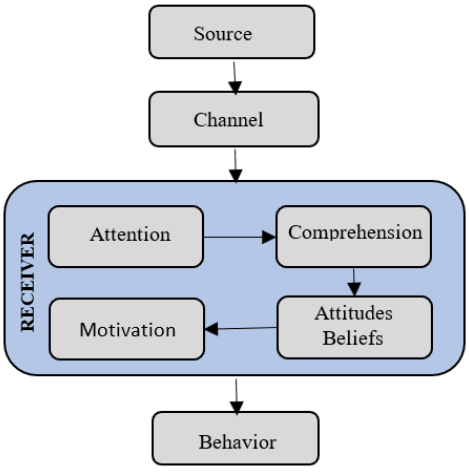

Fig. 1. Communication-Human Information Processing Model (C-HIP)

Channel: This is related to the way information is transmitted from the source to the receiver [76]. Ex-

isting solutions for privacy risk communication include push notifications or visualizations within an application [55, 59]. As noted in the related work section, notifications can be intrusive, and users may not be sufficiently motivated to access any in-app visualizations. To resolve these downsides, our "glanceable" approach suggests the use of a *background* channel that allows the information to be unobtrusively but pervasively transmitted (e.g., via the phone's screensaver or lock screen).

Attention: At this stage, the design and placement of the interface should be salient to ensure users are aware of what is being presented. Once they have noticed the communication, their attention needs to be maintained for a period long enough to gather information [76]. We argue that using a background channel gives users the opportunity to passively gather information periodically. This gives users a "background awareness" of the information-sharing practices of their smartphone apps. As such, any unexpected change in the presented visualization (e.g., a sudden surge of unauthorized data leakage) can serve as a meaningful cue to grab users' attention and inform them when immediate action may be needed.

Comprehension: After receiving the users' attention, the information needs to be presented in a way that aids comprehension. In this regard, there are difficult tradeoffs between concise and comprehensive notices which may have significant effects on users' understanding of what is being communicated [76], especially in the field of privacy. In fact, Nissenbaum argues that it is impossible to create sufficiently comprehensive privacy notices that are also concise enough to be understood [48]. Therefore, our designs account for varying levels of detail, and our study evaluates whether these different levels of granularity aid or hinder comprehension.

Attitudes and Beliefs: A notice that is understood by a user still runs the risk of being ineffective if it does not influence their attitude towards the notice [76]. Specifically, for users to act upon a notice, they must judge it to be accurate and important. Such attitudes can be influenced by persuasive content and presentation. Note, though, that the persuasiveness (and therefore the effectiveness) of a certain notice may depend on what people value: if the persuasive aspects of the notice do not align with the user's beliefs about what is important, they may not be motivated to act. Therefore, our data exposure visualizations vary in terms of the type of content on which the information design is centered, and our study evaluates which type of presentation elicits the strongest attitudes.

Motivation: Once users' attitudes towards a notice are favorable, the final hurdle is to motivate them to act. In this final stage, it is important to design notices that do not overwhelm nor underwhelm users, as this could affect their behavioral intention [76].

In the following section, we describe how we applied the design recommendations derived from the C-HIP model to the design of our data exposure visualizations. Note that for the evaluation presented in this paper, we focus on users' attention, comprehension, and attitudes/beliefs. The integration of the visualizations into a background channel (e.g., Android's screensaver or lock screen) and their effect on users' motivation and behavior requires a live deployment of the visualization and will, therefore, be covered in future work.

4 Study Setup

4.1 Prototype design

We adopted an iterative design process to create pervasive and unobtrusive visualizations of the data sharing practices of mobile applications. These prototypes were developed with the explicit goal of increasing users' attention, awareness and comprehension. Prior to conducting our user study with high-fidelity prototypes, we created low-fidelity wireframes during seven months of design iterations. Initial evaluations were conducted during bi-weekly lab meetings with an average of 15 HCI researchers and Ph.D. students, Skype presentations with collaborators and HCI experts, and through informal conversations with friends and acquaintances to garner feedback from a diverse subset of potential users. Earlier designs featured graphs and tables, which proved to be difficult for users to interpret, and were redesigned over time into the final versions (see Appendix).

Ultimately, we created eight designs, manipulated along two design dimensions to address our research questions (See Table: 1). Depending on the specific version, the visualizations presented the following information: what type of information was shared, which app shared the information, and when that information was shared. To reduce the complexity of the visualization, we limited it to display at most 6 apps and 6 data types (we envision that in the final implementation of our designs, users can choose which apps and data types to display, and that the number of "slots" in the design may dynamically be set to larger or smaller than 6,

depending on the screen size and interface magnification). Sample screens for each manipulated dimension are shown in Figures 2-5. Below we describe our motivation behind the design manipulations.

Presentation	Granularity			
App-Centric	Low	Moderate	High	Very High
Data-Centric	Low	Moderate	High	Very High

Table 1. Description of the design dimensions.

4.1.1 Four levels of Granularity

In his description of the C-HIP model, Wogalter et al. acknowledges the trade-off between interfaces that are comprehensive and concise, and their effect on user comprehension [76]. In line with this, we varied our designs on four levels of granularity: low, moderate, high, and very high. Designs with a low granularity simply communicated the total number of times data was shared (broken down by either app or data type; first row of Figure 2). Designs with a moderate granularity showed which app shared what type of information but offered no insight as to when the data sharing occurred (second row of Figure 2). At the high level of granularity, designs also showed roughly when each app last shared each type of information (represented by its position on a circle; third row of Figure 2). At the very high level of granularity, designs showed the total number of times data was shared and the exact time each app last shared the information to allow normative comparisons (using a clock-like visualization; last row of Figure 2).

While the designs with higher levels of granularity show more detailed information, their complexity may negatively impact the *Comprehension* step of the C-HIP model and reduce the "glanceability" of the design.

4.1.2 App-centric vs. Data-Centric Visualization

Wogalter et al. also acknowledges that users may only develop an actionable attitude towards a risk notice if its persuasive aspects align with the user's beliefs about what is important [76]. Prior research on visualization of data disclosures focused on the relationship between the data attribute that was disclosed and the entity disclosing it [1, 5, 22, 79]. Our glanceable designs were centered around either one of these two aspects: our designs are either app-centric (i.e., focused on who is sharing the data; examples: Skype, Dropbox, Facebook, Instagram, WhatsApp, and Google Plus) or data-centric (i.e., focused on what is being shared; examples: location, camera, contact, browser activity, microphone, and call log). Arguably, a user's beliefs may be more aligned with one of these two presentations, in which case that design will be more effective in the *Attitudes and Beliefs* step of the C-HIP model (which is a prerequisite for motivating users to act).

4.2 User Study Design

We used the eight designs as a probe in our semi-structured interviews to investigate the glanceability and understandability of our data exposure visualizations. To participate in our study, participants had to be over 18 years old and currently own an Android phone. We began recruitment in February 2018 and completed data collection in April 2018. Flyers were placed in public places around a large US university campus and additional participants were recruited from local coffee shops. Participation was incentivized with $5 Starbucks gift cards upon completion of the interview. Data were collected from a total of 15 participants (10 male, 5 female; age ranges: 18-24 (9), 25-34 (4), 35-44 (2); levels of education: high school degree only (3), some college but no degree (7), bachelor's degree (1), master's degree (4)). Table 2 summarizes participant demographics, their preferred structure of information (app-centric versus data-centric) and their preferred granularity (low, moderate, high, very high).

Before each interview, participants were asked to review a consent form. The researcher would then ask the participant to imagine having downloaded an app with an accompanying screensaver with a visualization that helps them track the data sharing practices of applications on their phone. The app would allow them to choose six apps and six data types that they would like highlighted in this visualization. In the study, we chose to hold the apps and data types constant between participants to avoid possible confounding effects and to reduce variability. Popular apps and data types were selected to increase the chance that participants would be familiar with them.

Each participant was shown two of the eight data exposure visualizations on their personal mobile phones: one app-centric and the other data-centric (in randomized order to control for ordering effects), at two different randomly selected levels of granularity. Each par-

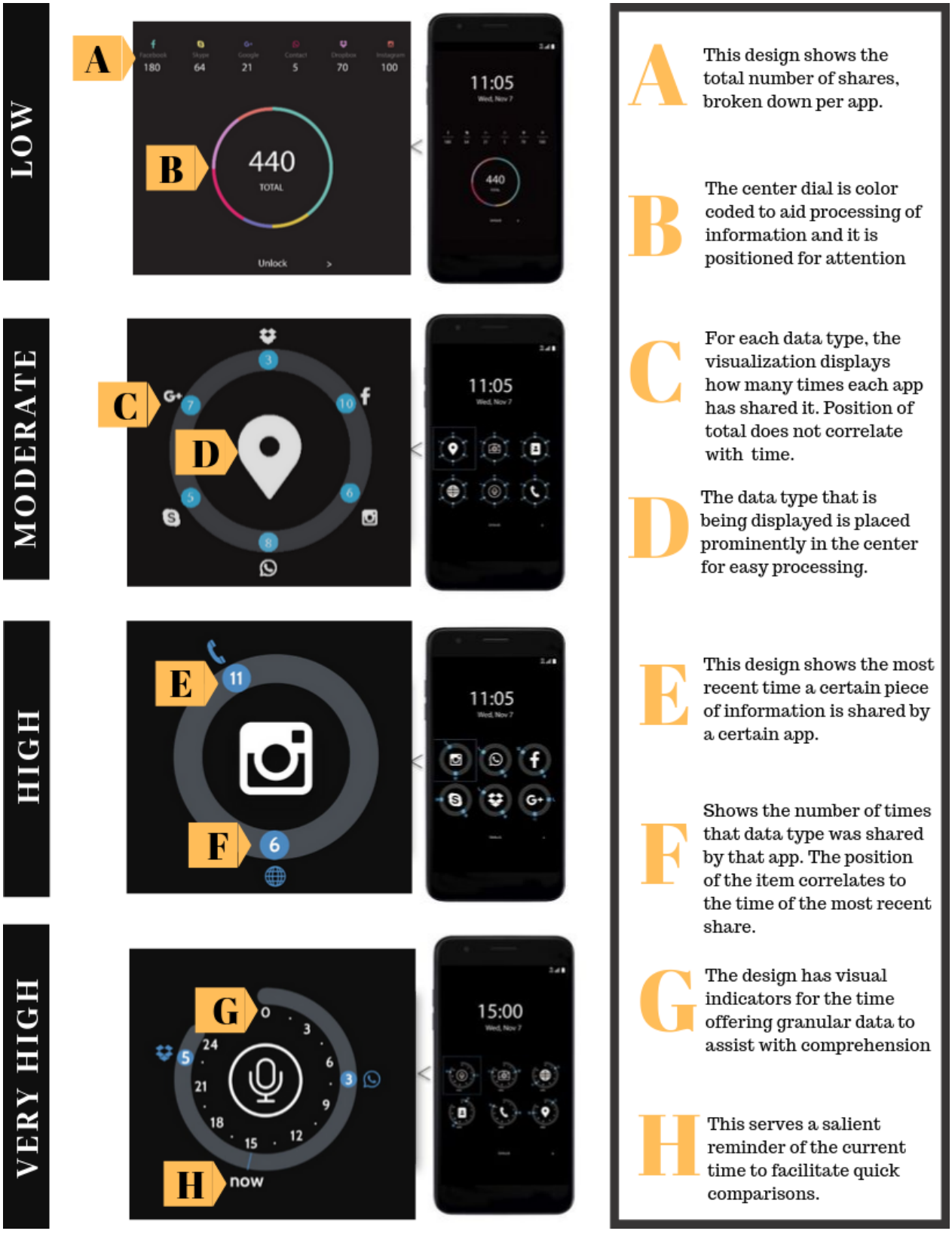

Fig. 2. Annotated Designs: Designs varied in their level of granularity (low, moderate, high, very high) and the presentation style (app-centric versus data centric) with a total of eight designs. For each design shown, there is an identical design with the same level of granularity but different presentation style. From top to bottom: "Low granularity, app-centric presentation", "Moderate granularity, data-centric presentation", "High granularity, app-centric presentation", "Very High granularity, data-centric presentation".

Participant	Gender	1st Design Shown	2nd Design Shown	Presentation preference	Granularity preference
P1	F	App-centric, Moderate	Data-centric, Low	A	M
P2	F	Data-centric ,Moderate	App-centric, Low	D	M
P3	M	Data-centric, Very High	App-centric ,High	A	H
P4	F	App-centric, Very High	Data-centric, High	D	H
P5	M	App-centric, Moderate	Data-centric, Low	A	M
P6	M	Datcentric, Moderate	App-centric, Low	D	M
P7*	F	App-centric, High	Data-centric, Very High	D	H
P8	F	Data-centric, High	App-centric,Very High	N	H
P9	M	App-centric, Moderate	Data-centric, Very High	A	M
P10	M	Data-centric, Low	App-centric, High	N	H
P11*	M	App-centric, Moderate	Data-centric, Very High	D	M
P12	M	App-centric, High	Data-centric, Low	A	H
P13	M	Data-centric, High	App-centric, Low	D	H
P14	M	Data-centric, Moderate	App-centric, High	A	H
P15	M	Data-centric, Very High	App-centric, Low	A	L

Table 2. Participant demographics. Presentation preference: [A]pp-centric, [D]ata type-centric, or [N]o preference/context-dependent. Granularity preference: [L]ow, [M]oderate, [H]igh, [V]ery [H]igh. *preferred a mix-match of the presentation and granularity shown.

ticipant (n=15) saw both presentation styles and each granularity type was seen by at least six participants. Given our focus on glanceability and comprehension rather than interaction, we presented the visualizations as static image-based prototypes. After the introduction, we used a 5-second test (5ST) to gather initial impressions and evaluate glanceability [27]. This usability technique is commonly used to evaluate users' understanding and first impressions of an interface. In the study, participants were shown each design for five seconds in the initial stage then a longer period for a prolonged inspection time. The researcher would then ask follow-up questions focused on participants' understanding of the visualizations as well as their comprehension of the data sharing situation depicted in the visualization.

After participants' first impression was captured, we gave them the opportunity to view each design for 2-5 minutes. We asked them what they thought the visualization was trying to display. Then we asked probing questions to gather additional insights about both their understanding of the visualization and their comprehension of the information depicted therein. Participants were asked about their thoughts about both designs if their preference did not come up naturally after inspecting both designs. Finally, we invited participants to share any final thoughts about the data exposure visualizations. Demographic information was collected at the end of the interview.

The average length of each interview was 32 minutes (min: 24, max: 40). Quotes and anecdotes from participants are presented throughout the paper with codes (P#) to protect their identities. This study was approved by the university's Institutional Review Board (IRB).

Theme	Description
Customization	**Having the ability to make custom changes to the interface to quickly gather information**
Visual Clues	**Having informative indicators for changes at a glance**
Data Flow	**Needing detail about the breakdown of information flow. Could be positive or negative**
Access	**Being able to easily access and view notifications about data sharing**
Accountability and Functionality	**Match between user behavior and what is being done with the data on the phone**
Transparency and Privacy Concerns	**Concerns about specific companies and their data sharing/data access patterns**

Table 3. Final Codebook: Themes are linked to the quotes in the results section.

4.3 Data Analysis Approach

Interviews were audio recorded and transcribed using Descript[1]. Participants' preferences were inferred from their responses concerning (1) whether they preferred the app-centric or data-centric designs, and (2) what level of granularity they preferred. The researchers worked collaboratively to achieve coding consensus and calculated IRR (Inter-rater reliability) for participants' preference using Fleiss' kappa, achieving a substantial level of reliability (k = 0.762) for the preferences summarized in Table 2. Open coding [68] was used to identify emerging themes from participants with three coders. The researchers conducted a thematic content analysis of the transcribed interview responses [28]. Themes were developed from the data as the analysis progressed, and once they were established the data was re-analyzed to uncover related data. After 12 participants, no new themes emerged. The researchers achieved saturation at 15 participants [19]. Below we describe the themes from our codebook.

5 Results

We organize our results by the initial research questions posed in the introduction and present the themes that emerged during our qualitative coding process within the framework of these research questions.

5.1 First Impressions of Glanceability (RQ1)

The glanceability of the designs were observed at two times: (a) during the five second test where first impressions were collected and, (b) at the second stage where participants were allowed a longer inspection time (~two minutes).

5.1.1 Supporting the Abstraction of Data

We found that participants' first impressions of the data exposure visualizations were positive: the visualizations attracted their attention as they intuitively attempted to decode their meaning (RQ1). A recurring principle that participants seemed to place emphasis on was the need to support the *abstraction of data* that would enable them to perceive and process information in a quick manner with minimal cognitive effort. As one participant noted:

"... If something is happening and I don't want it to happen, it is easy for me to catch it with a glance at this screensaver without searching deep to know what's going on." (Participant: P15, Theme: Visual Cues)

In general, most participants (12/15) were able to accurately describe the purpose of the privacy visualizations after viewing for five seconds. At the initial stage, participants did not focus on their perceptions of any particular apps or data types in a silo but rather how each component contributed to their understanding of the entire ecosystem of data being shared. One participant summarized the purpose of the visualization:

"To give me the information that says which app is sharing information [...] with which applications" (Participant: P9, Theme: Data Flow)

The design of the visualizations were salient enough to ensure participants were aware of what is being presented. For those instances where participants experienced misunderstandings about the purpose of the visualizations their opinions converged on similar points around granularity and glanceability which we discuss next.

5.1.2 Granularity versus Glanceability

Although most participants understood the purpose of the visualizations, the designs with the highest level of detail seemed to be overwhelming for participants at first glance. For instance, even after being able to explain the purpose of the visualization with a lower level of granular detail, when presented with a design of a higher level of granularity, P15 remarked:

"I'm totally clueless about what's going on on the screen. Judging from the design, I'm not able to figure out what is the purpose." (Participant: P15, Theme: Visual Cues)

This suggests that the purpose of the visualization is easier to understand upon first inspection with less granular designs compared to ones that offer more detail about potential data exposures.

Beyond the identification of the purpose of the visualizations, glanceability also contributed to how quickly and easily information from the visualizations were conveyed. In this regard, participants preferred the designs with lower levels of granularity since those designs were

1 Descript - Transcription and Audio Editing. Retrieved from http://descript.com

able to illustrate at a glance what the participant viewed as relevant. Referring to a design with low granularity versus a very high granularity design, P15 mentioned:

"The design is pretty good. It is not overflowing the screen with too much information. Making the information a little bit bigger, because the screen is kind of empty, would be easier to grab information with a simple glance." (Participant: P15, Theme: Visual Cues)

For a visualization to be perceived as glanceable, it was important that the essence of the information was conveyed in a simplified manner that gained participants attention, supported the abstraction of data, and could be processed in a meaningful way. Next, we reveal findings around the design components that influenced glanceability.

5.1.3 Glanceability Influencers

Participants' reactions helped in identifying a number of opportunities for how glanceability could be improved that would enable sustained glancing long-term. Across the different versions of designs, there were similarities among users in what they perceived to be helpful visual cues to improve glanceability.

Specifically, across the different designs, users highlighted the importance of being able to customize the order, color, and position of the icons in the visualization. They remarked that allowing more autonomy over customization would ultimately assist with memory association.

"I don't know if those are colors that you would assign yourself or if it just defaults but I think choosing gives more freedom" (Participant: P2, Theme: Customization)

Moreover, a recurring theme that emerged was around having visual components of the design that would assist in processing information in repetitive inspections of dynamic content. In particular, participants emphasized the importance of visual cues that would help to identify *novelty* in the privacy visualizations and ultimately assist in identifying patterns to plan a course of action. Contextually, this may be helpful when trying to gather an overview after long periods. For instance:

"... in the morning when you wake up and they're all zero or whatever they're at for the day. And then it says Facebook has shared your location five or six times it like turns green... or if it's gone down it could turn red or yellow whatever amazing color to read." (Participant: P1, Theme: Visual Cues)

At the same time, simplicity in design would be critical in reducing cognitive effort and aiding the processing of information. In line with this principle, P4 noted a point of contention with the privacy visualization with very high granular detail:

"I don't know why the 24-hour clock had to be split in terms of three...Maybe make it even numbers so I don't have to calculate and there's more beauty so it's more intuitive." (Participant: P4, Theme: Visual Cues)

Likewise, adopting a similar format to existing designs users would be familiar with assist in creating designs users would perceive as intuitive.

5.2 Feedback from Prolonged Inspection (RQ2)

After exploring the visualizations for a longer amount of time, participants were able to identify factors that motivated them to transition from glancing to exploration. Acting inline with the C-HIP model, the visualizations were successful in transitioning participants from attention to comprehension. At this stage, the visualizations served as a gateway to engagement and reaction to the information that has been conveyed.

5.2.1 Increased Awareness Inspires More Investigation

Glanceable privacy visualization offer insightful moments rather than a deep exploration of information. Thus, the designs act as cues for engagement. In this sense, participants suggested that the privacy visualizations adopt an approach where the information being conveyed raises questions versus providing answers.

"At 7:30 or 8:00-ish in the morning WhatsApp shared my location with three entities. Okay. The information has been shared. But I want to know with who though." (Participant: P8, Themes: Data Flow)

Upon further exploration, participants continued to raise questions as they attempted to understand the inter-relatedness of data, the entities they were being shared with, and how learning about this was related to their expectations from these entities.

"This can help you to know which apps are installing your information. You can know when Google Plus used your information and what kind information it was but maybe not why. And normally you never know but apps like WhatsApp could get your location from your last vacation then boom that's shared with the government."

(Participant: P11, Themes: Transparency and Privacy Concerns)

Participants displayed an understanding of data sharing practices and they were able to understand the flow of data but no one mentioned advertisers as a possible recipient of their data. Participants were mainly concerned about the data stored in particular apps being shared with other companies.

5.2.2 Less Can Be More

While the more granular visualizations were considered less glanceable, most (all but two) participants were able to comprehend even the visualizations with very high granularity. When given the opportunity to inspect the design for a longer period P15 expressed:

"After getting a closer look, now it is more clear. It is showing me the hours of the day. It provides more information than the first design. I know the time, when the data sharing happened, and what is requested by the apps, and app by app I can see what is going on." (Participant: P15, Theme: Data Flow)

High levels of granular detail about the data flow proved to be informative but overwhelming and not valuable. For instance, P7 said:

"I still don't understand the numbers in the middle and the 'now'... It almost seems like overkill and so just having the small numbers with a little icon would be informative enough." (Participant: P7, Theme: Visual Cues)

While participants were able to interpret the information in the more granular designs, some participants still opted for a design with less granular detail. Particularly, the information about the exact time the data sharing occurred was generally perceived as not valuable.

"I'm not sure I quite see the benefit of knowing exactly when like on the clock. Like the time function, I'm not sure I see the benefit of knowing when my information was shared. I mostly would care about if it was shared at all." (Participant: P11, Theme: Accountability)

As such, a common viewpoint was that the visualization should promote *actionable* insights that adequately informs without the need for deep exploration.

5.2.3 Advocating for Personalized Experiences

Participants intuitively predicted an inevitable increase in complexity if more apps/data types would be shown, and instead suggested the option to customize the order and type of content but still using our current layout. However, participants highlighted that the visualization should combine the principle of *simplistic design* with the benefits of a unique tailored experience.

"If there is a bunch of apps, the design might get too dense, but with this number of apps is not too bad. I would like to see options: first, the app with the most recently used. Second, the apps that used most permissions (data accessing) in the last hour or whatever." (Participant: P14, Theme: Customization)

Having an interface adapted to personal preferences was a reoccurring topic that participants expressed was fundamental in assisting with making progressive comparisons to identify the need for action.

5.2.4 Progressive Disclosure

Participants expressed the importance of eventually having access to a detailed breakdown of the data exposures by app, data type and time, but they argued that such information would not have to be immediately available. Upon investigating the visualization, one participant argued that the less granular design could serve as a proxy to a more detailed break-down within the app.

" ...it doesn't have to be on the screensaver. But I do want to be able to see the breakdown if I open the app or if I click the screensaver to open the app. Especially if you know Instagram wasn't 100 and the next time I check it it's like 150 and I'm like what is going on with Instagram right now?" (Participant: P1, Theme: Data Flow)

The visualizations were described as a tool that could help users determine if they needed to (a) investigate certain risky/unusual behaviors, (b) decide if they needed to pay more attention to a particular app or data type, or (c) take immediate action.

5.3 Data- vs. App-Centric Preferences (RQ3)

Interestingly, our participants were fairly split as to whether they preferred a data-centric versus app-centric presentation of the information: Six participants pre-

ferred the data-centric presentation, seven preferred the app-centric presentation, and two participants did not have a preference. Yet, we uncovered unique characteristics of our participants based on their preference and describe those characteristics in more detail below.

5.3.1 Data-centric: General Risk Assessment and Awareness of Data Sharing

The six participants who preferred the data-centric visualizations expressed stronger concern for awareness of data sharing, which permissions were being accessed, and specific types of data being shared without their knowledge.

"I don't know if I need to know how much data is being shared, just if that information is being shared." (Participant: P13, Theme: Accountability)

For some, using data-centric designs provided the opportunity to assess potential risks to privacy at a higher level while allowing further inspection to aid the decision-making process.

"I think my brain personally works best with knowing what type of information is being shared first, which is why I like these simple icons and then I can be like, oh well location, who wanted my location and then look further and see that it was WhatsApp, Dropbox or something" (Participant: P8, Theme: Accountability)

For these participants, having clarity of the type of data being shared and being able to monitor the flow of sharing is perceived as more valuable in helping them to gauge the level of risk.

5.3.2 App-centric: Distrust in Companies and the Need for Transparency

Participants who preferred app-centric designs (7 of 15) routinely expressed more value in knowing who shared their information and being informed about the data sharing activities regardless of the data type.

"It's pretty cool like to see how many times stuff has been shared and makes you know who it is shared with. Makes you want to know who has access to my data" (Participant: P12, Theme: Transparency)

Knowing the frequency of data sharing as an indicator of a company's data sharing habits was often mentioned.

"It makes it [clearer] what each app is using. It ties my thinking to what the app is doing rather than these permissions are being used by these apps." (Participant: P3, Theme: Transparency)

"Apps like Google, Facebook and Instagram...it just kind of adds up and shows you how much like the amount of data" (Participant: P5, Theme: Transparency)

Unlike participants who preferred a data-centric presentation, those with an app-centric preference were more motivated to investigate which entities showed activities that weren't aligned with their expectation of data usage and sharing.

"I wouldn't have thought that Dropbox is accessing microphone and camera. So that is interesting. I don't think using that is necessary for sharing files. Other things like using contacts for social media apps like Facebook, that makes sense and I understand why that's happening. There is obvious data sharing that I understand why it is happening, and there are others that I would change my settings." (Participant: P14, Theme: Transparency)

Some participants in this group also showed stronger feelings towards surveillance and needing privacy.

"You don't know they're like using it in the background, it kind of like tracks you where you are. So, it's kind of scary." (Participant: P3, Theme: Transparency)

"Um, my phone does this? This kind of scares me. I want to turn it [all sharing] off and go in a cabin where no one can find me." (Participant: P1, Theme: Transparency)

Many participants with app-centric preferences expressed that the visualizations reminded them how little control they had over their personal information and this fundamentally changed how they saw apps and the companies that own them.

5.3.3 A Bit of Both Worlds

Finally, participants (2/15) who had no preference seemed to demonstrate some concern for general awareness as well as distrust in certain specific companies.

"It depends on what you care about because if you're concerned about a specific app sharing your information I want to track all of the different things that that specific app was doing but I know for the other apps I would be like I'm concerned about my information being shared at all." (Participant: P8, Theme: Transparency)

For these participants, the presentation style (data-centric or app-centric) was content dependent. Participants lamented that all apps are created equally and

their intrinsic level of concern may vary depending on their existing suspicions of companies or distrust with mobile platforms and specific data types in general.

6 Discussion

In this section, we revisit our research questions and provide supporting results for each. We draw comparisons between our findings and prior work and identify contributions. Subsequently, we list potential implications for privacy research, design, and future work. We end this section with a discussion of the limitations of our methodology.

6.1 Transition from Glancing to Understanding

In line with the C-HIP model, we tested the effect of glaceability as a prerequisite of the Attention step, the effect of granularity on the Comprehension step, and the effect of presentation style that matches users' attitudes on the Motivation step. Our analysis suggests that the use of glanceable data exposure visualizations supports users' engagement with and understanding of mobile app data sharing practices. When shown for the initial five seconds, the visualizations maintained participants' attention as they intuitively attempted to decode the meaning of the presented information. Most participants were able to accurately describe the purpose of the visualizations within those first five seconds, although designs with the highest level of granularity (see Figure 2) proved to contain too much detail to process. As a result, our most detail-oriented designs (i.e. with very high granularity) were perceived to be too complicated for participants even when they were presented with a prolonged time for inspection (i.e., after participants had been given the time to fully interpret another design), let alone at first glance. Considering these designs had additional temporal information, it is possible that may have contributed to an overload of information that could be consumed at a glance. Being able to identify nuances in information influences users' ability to distinguish critical information, and it ultimately impacts the glanceability of the interface (RQ1).

Moreover, our findings suggest that glanceability acts as a mediator between the level of granularity of the interface and user comprehension (granularity → glanceability → comprehension). We saw that after expressing a detailed understanding of the information flow (i.e., knowing who shared what type of information and when), some participants consciously weighed the advantages of knowing this level of detail with the reduced glanceability of the interface. Particularly, aspects that were less glanceable (e.g. the time of disclosure) were also deemed less useful upon further inspection. Conversely, this suggests that the factors that are more likely to capture the attention of users influence what they deem relevant for their understanding. This is in line with the flow of the C-HIP model [76] (attention → comprehension) and helps us to better understand (a) the trade-offs users are willing to make and (b) the aspects of the design that contribute to users' understanding of data exposures.

6.2 Conceptualizing Privacy: Relational versus Content

Participants varied in their preference for a data-centric or app-centric presentation of the information: a group of six participants leaned towards a data-centric presentation, a group of seven leaned towards an app-centric presentation, while the two remaining users displayed characteristics of both groups (RQ2). Within each group, there were distinct cognitive differences in what was perceived to be valuable and how to process privacy-related risks. In line with the Communication Privacy Management theory, participants' views on privacy decision-making with other parties were dependent on their personal disclosure preferences and whether they felt that the app was an appropriate co-owner of this information [53]. Participants in both groups expressed a sense of ownership of their data, but they employed different strategies to maintain their privacy.

Specifically, participants who preferred data-centric information placed stronger boundaries around specific *types of content* regardless of the recipient. Consequently, monitoring the flow of that data was more important, and greater emphasis was placed on having content-based control. Conversely, participants who preferred app-centric designs tended to place boundaries around recipients and were more concerned about entities violating their trust as they were perceived to be co-owners of their personal information. Therefore, maintaining privacy for those participants is more *relational* and it would require more effort to not only monitor but to investigate the data sharing activities of applications to ensure that disclosure is aligned with their expectations.

These diverging conceptualizations of privacy boundary regulation are similar to prior privacy research in networked privacy [31, 71], which shows that users agree to sacrifice some of their privacy for functionality (e.g. using smartphone apps) as long as their privacy boundaries are respected. Our work is unique in demonstrating that personal differences in the importance of certain privacy boundaries are reflected in users' preferences for the design of privacy-enhancing technologies, such as our data exposure visualizations.

Moreover, contextual considerations may have implications for the preferred structure of the information as well, as some of our participants reported that both structures of the information would be useful, depending on the situation: users may prefer a data-centric approach to achieve general awareness, but an app-centric approach to monitor newly installed apps they do not yet trust.

These personal and contextual variations are key to the final step in the C-HIP model: motivation—arguably, a data exposure visualization is more likely to be able to motivate the user to take action if the structure of its information matches the user's personal and contextual preferences.

6.3 Design Implications

Based on our findings, we identify insights and recommendations for HCI and privacy researchers, designers, and developers that may benefit end-users and companies.

6.3.1 Develop Attention-Sustaining but Meaningful Notifications

Facilitating users' ability to abstract data is a fundamental quality in glanceable interfaces and designs [26, 46, 60]. Our findings suggest this is an important design aspect for glanceable privacy visualizations to maintain the Attention and Comprehension steps of the C-HIP model. Attention-sustaining designs would adopt a balance between *simplicity* and *novelty* - helping users to quickly identify new risks while avoiding notification overload.

While our findings suggest that users derive value from using our glanceable visualizations to improve their privacy awareness, this does not mean that using a background channel to present visualizations is a panacea for privacy notifications. For one, we would have to implement and test these visualizations in an actual smartphone (e.g., as part of the Android lock screen or screensaver) to validate the use of this channel. Moreover, we encourage researchers and designers to explore different channels for privacy risk communication, or new ways in which existing communication channels could be made more glanceable and meaningful for users. Moving forward, researchers could consider supporting granular information in risk communications with appropriate visual cues to improve glanceability while providing alternatives that would best align with what users value. Future studies could also consider other structural dimensions such as the expectedness of data flows [56]. Similarly, designers should consider the limited real estate of their channel (i.e., screensavers or push notifications), as a lack of readability could have serious implications, such as not recognizing risks.

6.3.2 A Case for Personalized and Adaptive Designs

Another opportunity for design would be adopting a user-tailored approach. Our results suggest that personal and contextual factors determine what is the most meaningful method for displaying information in a data exposure visualization. In future work, researchers could consider a user-tailored approach to privacy, e.g. by first measuring the user's privacy boundary regulation preferences, then using the measurements to adapt the structure of the presented information to the predicted privacy preferences [8, 36, 75]. Personalized and adaptive designs, may better align with users' need for customization to promote insightful "aha!" moments and maintain the Comprehension stage in the C-HIP model. However, this approach should be implemented with caution, as using personalized recommendations may itself have privacy implications.

6.3.3 Consider Deployment Models

One of the final stages of the C-HIP model involves motivating users to act. Stakeholders should consider that privacy is usually a secondary activity of using technologies [42]. As such, dedicated applications that reveal privacy leaks may only appeal to users who are already concerned about their privacy. Similarly, researchers should consider the impact of chosen channel on users' motivation. Having a background channel like a screensaver may by helpful in improving the salience of privacy visualizations. However, if widely available for

download, it's possible that apps that require access to unfamiliar channels may be perceived as suspicious to users. As such, solutions released from platform developers may be seen as more trustworthy and stakeholders in this domain are encouraged to test and incorporate new tools.

6.4 Limitations and Future Work

Our methodology allowed us to gain a deeper understanding of users' reactions to our data exposure visualizations. The interview setting allowed us to gather first-hand insights into the factors that shaped users' perspectives, however, there are potential limitations to consider. Providing context has the potential to influence behavior and it is difficult to extrapolate whether a static visualization presented in such an interview setting will be perceived the same way when integrated into the user's actual screensaver or lock screen. Indeed, privacy researchers have found evidence that supports the 'privacy paradox' [49] which implies that users may find new privacy tools useful in controlled settings, but their reaction and use may be significantly diminished when said tool is used in practice.

To minimize potential carryover effects, each participant was exposed to both of the presentation styles but only two of the four levels of granularity. Future studies may want to consider methodological options that would allow participants to investigate different combinations of granularity to further explore users' preferences.

We also held the information shown in the visualizations constant between participants, rather than tailoring the visualizations to participants' actual data exposure. This avoided the need for a full field trial, and enabled the researchers to reduce variability in user responses, but it could be argued that participants' behavior may have been affected by the fact that the provided artificial information did not reflect participants' real-life data exposure. The results from this interview study are relevant nonetheless, as they inform and contextualize future live field trials with a real "privacy screensaver". In these trials, we will be able to observe users' actual adoption and use of our visualizations, and whether they inspire longer-term changes in behavior and privacy decision-making. These field trials would not, however, give us the detailed insights regarding granularity, glanceability and information structure that we gained from the current study.

One area where the generalizability of our results is limited, is in the analysis of potential "interaction effects" between the data type and the level of granularity of each visualization. Our results provide separate, qualitative insights regarding data type and granularity; in a field trial with a sufficiently large sample, one could potentially investigate whether the optimal level of granularity is the same or different for each data type. The primary aim of our current study was to identify qualitative differences between data types and levels of granularity, but we certainly suggest that future work involves confirmatory studies to quantitatively investigate potential interaction effects between these two parameters. As we present our findings based on the data of our participants, it is important to keep in mind that the demographics of our sample were skewed towards young and educated US residents. Caine reported that this sample size is not uncommon for qualitative HCI research [13]. Although we achieved saturation, there may be limitations in interpreting the findings on a larger scale, as our sample size may limit generalizability.

Looking forward, future work should also consider factors that would affect deployment such as the trade off between information leakage detection accuracy in a real time and the impact on users' battery life. Existing tools such as Antmonitor shows that 99% of leaked packets were mapped correctly to the installed applications, while only 3% of the battery overhead is been measured in average mobile usage [40]. PrivacyProxy is another application that proves the perceived impact on battery life by users when using the app is low[67].

7 Conclusion

Despite the large body of existing work on privacy-enhancing technologies, the problem of "helping users understand which applications share what type personal information and when" still provides a wide range of questions for researchers to explore. In this study, we introduced glanceable data exposure visualizations as a novel approach to this problem and offered insights into how such visualizations are perceived by users. Through our design probes and interviews with 15 participants, we provide insights into how users process information about privacy risks, and how the design aspects of data exposure visualizations aid or hinder the comprehension of those visualizations. We find that users' preferences for app-centric or data-centric information depend on whether they value relational or content privacy bound-

aries. Furthermore, we find that providing less information may at times be more useful, as aspects such as the granularity of the details impact users' understanding of data exposure and the glanceability of the interface.

Our findings suggest opportunities for the design of new mechanisms for the visualization of privacy that show the "right" amount of information while offering users the opportunity to further explore and investigate potential privacy risks. Finally, our work opens up lines for further studies to explore pervasive but unobtrusive privacy visualizations, and work towards improving privacy awareness for smartphone users.

8 Acknowledgements

We would like to acknowledge and thank Dr. Kelly Caine for her guidance and feedback throughout the duration of this project. We also thank the members of the Clemson University HATLab (Humans and Technology Lab) for offering critiques about the methodology and early designs. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] "Mozilla: Lightbeam." [Online]. Available: https://addons.mozilla.org/en-US/firefox/addon/lightbeam/

[2] A. Acquisti, "Nudging privacy: The behavioral economics of personal information," *IEEE security & privacy*, vol. 7, no. 6, pp. 82–85, 2009.

[3] A. Adams and M. A. Sasse, "Users are not the enemy," *Communications of the ACM*, vol. 42, no. 12, pp. 40–46, 1999. [Online]. Available: http://dl.acm.org/citation.cfm?id=322806

[4] Y. Agarwal and M. Hall, "Protectmyprivacy: detecting and mitigating privacy leaks on ios devices using crowdsourcing," in *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. ACM, 2013, pp. 97–110.

[5] J. Angulo, S. Fischer-Hübner, T. Pulls, and E. Wästlund, "Usable transparency with the data track: a tool for visualizing data disclosures," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2015, pp. 1803–1808.

[6] A. Azfar, K.-K. R. Choo, and L. Liu, "Forensic taxonomy of android productivity apps," *Multimedia Tools and Applications*, vol. 76, no. 3, pp. 3313–3341, 2017.

[7] M. Backes, S. Bugiel, and E. Derr, "Reliable third-party library detection in android and its security applications," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 356–367.

[8] P. Bahirat, Y. He, A. Menon, and B. Knijnenburg, "A Data-Driven Approach to Developing IoT Privacy-Setting Interfaces," in *23rd International Conference on Intelligent User Interfaces*. ACM, 2018, pp. 165–176.

[9] K. Benton, L. J. Camp, and C. Small, "OpenFlow vulnerability assessment," in *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*. ACM, 2013, pp. 151–152.

[10] R. Böhme and J. Grossklags, "The security cost of cheap user interaction," in *Proceedings of the 2011 workshop on New security paradigms workshop*. ACM, 2011, pp. 67–82. [Online]. Available: http://dl.acm.org/citation.cfm?id=2073284

[11] R. Binns, U. Lyngs, M. Van Kleek, J. Zhao, T. Libert, and N. Shadbolt, "Third Party Tracking in the Mobile Ecosystem," *arXiv preprint arXiv:1804.03603*, 2018.

[12] D. M. Boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," *Journal of computer-mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.

[13] K. Caine, "Local Standards for Sample Size at CHI," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 981–992, event-place: San Jose, California, USA. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858498

[14] P. R. Center, "Mobile Fact Sheet," *Pew Research Center: Internet, Science & Tech*, Feb. 2018. [Online]. Available: http://www.pewinternet.org/fact-sheet/mobile/

[15] L. Cerejo, "Glanceability & The Glanceable User Experience," Jul. 2013. [Online]. Available: https://www.capgemini.com/2013/07/glanceability-the-glanceable-user-experience/

[16] T. Chen, I. Ullah, M. A. Kaafar, and R. Boreli, "Information Leakage Through Mobile Analytics Services," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*, ser. HotMobile '14. New York, NY, USA: ACM, 2014, pp. 15:1–15:6. [Online]. Available: http://doi.acm.org/10.1145/2565585.2565593

[17] P. H. Chia, Y. Yamamoto, and N. Asokan, "Is this app safe?: a large scale study on application permissions and risk signals," in *Proceedings of the 21st international conference on World Wide Web*. ACM, 2012, pp. 311–320. [Online]. Available: http://dl.acm.org/citation.cfm?id=2187879

[18] R. Compañó and W. Lusoli, "The policy maker's anguish: Regulating personal data behavior between paradoxes and dilemmas," in *Economics of Information Security and Privacy*. Springer, 2010, pp. 169–185.

[19] I. Dey, *Qualitative data analysis: A user friendly guide for social scientists*. Routledge, 2003.

[20] W. Enck, P. Gilbert, S. Han, V. Tendulkar, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth, "TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones," *ACM Transactions on Computer Systems (TOCS)*, vol. 32, no. 2, p. 5, 2014.

[21] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner, "Android permissions: User attention, comprehension, and behavior," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*. ACM, 2012, p. 3. [Online]. Available: http://dl.acm.org/citation.cfm?id=

2335360

[22] S. Fischer-Hübner, J. Angulo, F. Karegar, and T. Pulls, "Transparency, Privacy and Trust–Technology for Tracking and Controlling My Data Disclosures: Does This Work?" in *IFIP International Conference on Trust Management*. Springer, 2016, pp. 3–14.

[23] H. Fu, Y. Yang, N. Shingte, J. Lindqvist, and M. Gruteser, "A field study of run-time location access disclosures on android smartphones," *Proc. USEC*, vol. 14, 2014.

[24] R. Gouveia, E. Karapanos, and M. Hassenzahl, "How do we engage with activity trackers?: a longitudinal study of habito," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 1305–1316.

[25] R. Gouveia, F. Pereira, A. Caraban, S. A. Munson, and E. Karapanos, "You have 5 seconds: designing glanceable feedback for physical activity trackers," in *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 2015, pp. 643–647.

[26] R. Gouveia, F. Pereira, E. Karapanos, S. A. Munson, and M. Hassenzahl, "Exploring the design space of glanceable feedback for physical activity trackers," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 144–155.

[27] G. Gronier, "Measuring the First Impression: Testing the Validity of the 5 Second Test," *J. Usability Studies*, vol. 12, no. 1, pp. 8–25, Nov. 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=3040226.3040228

[28] G. Guest, K. MacQueen, and E. Namey, *Applied Thematic Analysis*. SAGE Publications, Inc., 2012. [Online]. Available: http://methods.sagepub.com/book/applied-thematic-analysis

[29] R. Herbster, S. DellaTorre, P. Druschel, and B. Bhattacharjee, "Privacy Capsules: Preventing Information Leaks by Mobile Apps," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '16. New York, NY, USA: ACM, 2016, pp. 399–411. [Online]. Available: http://doi.acm.org/10.1145/2906388.2906409

[30] Q. Ismail, T. Ahmed, K. Caine, A. Kapadia, and M. Reiter, "To permit or not to permit, that is the usability question: Crowdsourcing mobile apps' privacy permission settings," *Proceedings on Privacy Enhancing Technologies*, vol. 2017, no. 4, pp. 119–137, 2017.

[31] P. Karr-Wisniewski, D. Wilson, and H. Richter-Lipford, "A new social order: Mechanisms for social network site boundary regulation," in *Americas Conference on Information Systems, AMCIS*, 2011. [Online]. Available: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1141&context=amcis2011_submissions

[32] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder, "A "Nutrition Label" for Privacy," in *Proceedings of the 5th Symposium on Usable Privacy and Security*, ser. SOUPS '09. New York, NY, USA: ACM, 2009, pp. 4:1–4:12. [Online]. Available: http://doi.acm.org/10.1145/1572532.1572538

[33] P. G. Kelley, S. Consolvo, L. F. Cranor, J. Jung, N. Sadeh, and D. Wetherall, "A Conundrum of Permissions: Installing Applications on an Android Smartphone," in *Financial Cryptography and Data Security*, ser. Lecture Notes in Computer Science, J. Blyth, S. Dietrich, and L. J. Camp, Eds., vol. 7398. Springer Berlin Heidelberg, 2012, pp. 68–79.

[34] P. G. Kelley, L. F. Cranor, and N. Sadeh, "Privacy As Part of the App Decision-making Process," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: ACM, 2013, pp. 3393–3402. [Online]. Available: http://doi.acm.org/10.1145/2470654.2466466

[35] P. Klasnja, S. Consolvo, D. W. McDonald, J. A. Landay, and W. Pratt, "Using mobile & personal sensing technologies to support health behavior change in everyday life: lessons learned," in *AMIA Annual Symposium Proceedings*, vol. 2009. American Medical Informatics Association, 2009, p. 338.

[36] B. P. Knijnenburg, "Privacy? I Can't Even! Making a Case for User-Tailored Privacy," *IEEE Security & Privacy*, vol. 15, no. 4, pp. 62–67, 2017.

[37] B. P. Knijnenburg and A. Kobsa, "Making decisions about privacy: information disclosure in context-aware recommender systems," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 3, no. 3, p. 20, 2013.

[38] H. D. Laswell, "The structure and function of communication in society," *The communication of ideas*, 1948.

[39] A. Le, J. Varmarken, S. Langhoff, A. Shuba, M. Gjoka, and A. Markopoulou, "AntMonitor: A system for monitoring from mobile devices," in *Proceedings of the 2015 ACM SIGCOMM Workshop on Crowdsourcing and Crowdsharing of Big (Internet) Data*. ACM, 2015, pp. 15–20.

[40] ——, "Antmonitor: A system for monitoring from mobile devices," in *Proceedings of the 2015 ACM SIGCOMM Workshop on Crowdsourcing and Crowdsharing of Big (Internet) Data*. ACM, 2015, pp. 15–20.

[41] J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang, "Expectation and Purpose: Understanding Users' Mental Models of Mobile App Privacy Through Crowdsourcing," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 501–510. [Online]. Available: http://doi.acm.org/10.1145/2370216.2370290

[42] Y.-H. Lin, C.-H. Fang, and C.-L. Hsu, "Determining Uses and Gratifications for Mobile Phone Apps," in *Future Information Technology*, ser. Lecture Notes in Electrical Engineering, J. J. J. H. Park, Y. Pan, C.-S. Kim, and Y. Yang, Eds. Springer Berlin Heidelberg, 2014, pp. 661–668.

[43] H. R. Lipford, A. Besmer, and J. Watson, "Understanding Privacy Settings in Facebook with an Audience View," in *Proceedings of the 1st Conference on Usability, Psychology, and Security*, ser. UPSEC'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 2:1–2:8. [Online]. Available: http://dl.acm.org/citation.cfm?id=1387649.1387651

[44] B. Liu, M. S. Andersen, F. Schaub, H. Almuhimedi, S. A. Zhang, N. Sadeh, Y. Agarwal, and A. Acquisti, "Follow my recommendations: A personalized privacy assistant for mobile app permissions," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, 2016, pp. 27–41.

[45] B. Liu, J. Lin, and N. Sadeh, "Reconciling Mobile App Privacy and Usability on Smartphones: Could User Privacy Profiles Help?" in *Proceedings of the 23rd International Conference on World Wide Web*, ser. WWW '14. New York,

NY, USA: ACM, 2014, pp. 201–212. [Online]. Available: http://doi.acm.org/10.1145/2566486.2568035

[46] T. Matthews, D. Blais, A. Shick, J. Mankoff, J. Forlizzi, S. Rohrbach, and R. Klatzky, "Evaluating glanceable visuals for multitasking," Technical Report EECS-2006-173. UC Berkeley, Tech. Rep., 2006.

[47] W. Nayam, A. Laolee, L. Charoenwatana, and K. Sripanidkulchai, "An Analysis of Mobile Application Network Behavior," in *Proceedings of the 12th Asian Internet Engineering Conference*, ser. AINTEC '16. New York, NY, USA: ACM, 2016, pp. 9–16. [Online]. Available: http://doi.acm.org/10.1145/3012695.3012697

[48] H. Nissenbaum, "A contextual approach to privacy online," *Daedalus*, vol. 140, no. 4, pp. 32–48, 2011.

[49] P. A. Norberg, D. R. Horne, and D. A. Horne, "The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors," *Journal of Consumer Affairs*, vol. 41, no. 1, pp. 100–126, Jun. 2007. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/j.1745-6606.2006.00070.x/abstract

[50] A. Oglaza, R. Laborde, A. Benzekri, and F. Barrère, "A Recommender-Based System for Assisting Non-technical Users in Managing Android Permissions," in *2016 11th International Conference on Availability, Reliability and Security (ARES)*, Aug. 2016, pp. 1–9.

[51] A. Oulasvirta, T. Rattenbury, L. Ma, and E. Raita, "Habits make smartphone use more pervasive," *Personal and Ubiquitous Computing*, vol. 16, no. 1, pp. 105–114, 2012.

[52] X. W. Page, *Factors that Influence Adoption and Use of Location-Sharing Social Media*. University of California, Irvine, 2014. [Online]. Available: http://search.proquest.com/openview/6b7ed26f5c311b4d1691b2d966cf3873/1?pq-origsite=gscholar&cbl=18750&diss=y

[53] S. Petronio, *Boundaries of Privacy: Dialectics of Disclosure*. SUNY Press, Feb. 2012, google-Books-ID: 8v89W_oJQ0wC.

[54] M. Pielot, A. Vradi, and S. Park, "Dismissed!: a detailed exploration of how mobile phone users handle push notifications," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 2018, p. 3.

[55] P. Rajivan and J. Camp, "Influence of Privacy Attitude and Privacy Cue Framing on Android App Choices," in *Symposium on Usable Privacy and Security (SOUPS)*, 2016. [Online]. Available: https://www.usenix.org/system/files/conference/soups2016/wpi16_paper-rajivan.pdf

[56] A. Rao, F. Schaub, N. Sadeh, A. Acquisti, and R. Kang, "Expecting the unexpected: Understanding mismatched privacy expectations online," in *Symposium on Usable Privacy and Security (SOUPS)*, vol. 4, 2016, p. 2.

[57] A. Razaghpanah, N. Vallina-Rodriguez, S. Sundaresan, C. Kreibich, P. Gill, M. Allman, and V. Paxson, "Haystack: In situ mobile traffic analysis in user space," *ArXiv e-prints*, 2015.

[58] J. Ren, M. Lindorfer, D. J. Dubois, A. Rao, D. Choffnes, and N. Vallina-Rodriguez, "Bug Fixes, Improvements,... and Privacy Leaks," 2018.

[59] J. Ren, A. Rao, M. Lindorfer, A. Legout, and D. Choffnes, "ReCon: Revealing and Controlling PII Leaks in Mobile Network Traffic," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '16. New York, NY, USA: ACM, 2016, pp. 361–374. [Online]. Available: http://doi.acm.org/10.1145/2906388.2906392

[60] Y. Rogers, W. R. Hazlewood, P. Marshall, N. Dalton, and S. Hertrich, "Ambient influence: Can twinkly lights lure and abstract representations trigger behavioral change?" in *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 2010, pp. 261–270.

[61] R. Roshandel and R. Tyler, "User-centric Monitoring of Sensitive Information Access in Android Applications," in *Proceedings of the Second ACM International Conference on Mobile Software Engineering and Systems*, ser. MOBILESoft '15. Piscataway, NJ, USA: IEEE Press, 2015, pp. 144–145. [Online]. Available: http://dl.acm.org/citation.cfm?id=2825041.2825076

[62] G. L. Scoccia, I. Malavolta, M. Autili, A. Di Salle, and P. Inverardi, "User-centric Android Flexible Permissions," in *Proceedings of the 39th International Conference on Software Engineering Companion*, ser. ICSE-C '17. Piscataway, NJ, USA: IEEE Press, 2017, pp. 365–367. [Online]. Available: https://doi.org/10.1109/ICSE-C.2017.84

[63] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.

[64] I. Shklovski, S. D. Mainwaring, H. H. Skúladóttir, and H. Borgthorsson, "Leakiness and creepiness in app space: Perceptions of privacy and mobile app use," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 2347–2356.

[65] H. J. Smith, S. J. Milberg, and S. J. Burke, "Information privacy: measuring individuals' concerns about organizational practices," *MIS quarterly*, pp. 167–196, 1996. [Online]. Available: http://www.jstor.org/stable/249477

[66] G. Srivastava, S. Chitkara, K. Ku, S. K. Sahoo, M. Fredrikson, J. Hong, and Y. Agarwal, "PrivacyProxy: Leveraging Crowdsourcing and In Situ Traffic Analysis to Detect and Mitigate Information Leakage," *arXiv preprint arXiv:1708.06384*, 2017.

[67] G. Srivastava, S. Chitkara, K. Ku, S. K. Sahoo, M. Fredrikson, J. I. Hong, and Y. Agarwal, "Privacyproxy: Leveraging crowdsourcing and in situ traffic analysis to detect and mitigate information leakage," *ArXiv*, vol. abs/1708.06384, 2017.

[68] A. Strauss and J. Corbin, *Basics of qualitative research: Procedures and techniques for developing grounded theory*. Thousand Oaks, CA: Sage, 1998.

[69] J. A. Tran, K. S. Yang, K. Davis, and A. Hiniker, "Modeling the engagement-disengagement cycle of compulsive phone use," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 2019, p. 312.

[70] M. Van Kleek, R. Binns, J. Zhao, A. Slack, S. Lee, D. Ottewell, and N. Shadbolt, "X-Ray Refine: Supporting the Exploration and Refinement of Information Exposure Resulting from Smartphone Apps," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: ACM, 2018, pp. 393:1–393:13. [Online]. Available: http://doi.acm.org/10.1145/3173574.3173967

[71] J. Vitak, S. Blasiola, S. Patil, and E. Litt, "Balancing audience and privacy tensions on social network sites:

Strategies of highly engaged users," *International Journal of Communication*, vol. 9, p. 20, 2015.

[72] Y. Wang, P. G. Leon, K. Scott, X. Chen, A. Acquisti, and L. F. Cranor, "Privacy Nudges for Social Media: An Exploratory Facebook Study," in *Proceedings of the 22Nd International Conference on World Wide Web*, ser. WWW '13 Companion. New York, NY, USA: ACM, 2013, pp. 763–770. [Online]. Available: http://doi.acm.org/10.1145/2487788.2488038

[73] D. Wetherall, D. R. Choffnes, B. Greenstein, S. Han, P. Hornyack, J. Jung, S. E. Schechter, and X. S. Wang, "Privacy Revelations for Web and Mobile Apps." in *HotOS*, 2011. [Online]. Available: http://static.usenix.org/event/hotos/tech/final_files/Wetherall.pdf

[74] P. Wijesekera, A. Baokar, A. Hosseini, S. Egelman, D. Wagner, and K. Beznosov, "Android permissions remystified: A field study on contextual integrity," in *24th USENIX Security Symposium (USENIX Security 15)*, 2015, pp. 499–514.

[75] P. J. Wisniewski, B. P. Knijnenburg, and H. R. Lipford, "Making privacy personal: Profiling social network users to inform privacy education and nudging," *International Journal of Human-Computer Studies*, vol. 98, pp. 95–108, Feb. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1071581916301185

[76] M. S. Wogalter, D. M. DeJoy, and K. R. Laughery, "Organizing theoretical framework: a consolidated communication-human information processing (C-HIP) model," *Warnings and risk communication*, pp. 15–23, 1999. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=6SuktVZ5UoAC&oi=fnd&pg=PA13&dq=Wogatler+C+hip&ots=LBRr4ZDRiE&sig=aUlatUT9dtRK09tGa16g-M0U8t8

[77] H. Xu, H.-H. Teo, B. C. Tan, and R. Agarwal, "Research note-effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: a study of location-based services," *Information Systems Research*, vol. 23, no. 4, pp. 1342–1363, 2012. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/isre.1120.0416

[78] J. Zang, K. Dummit, J. Graves, P. Lisker, and L. Sweeney, "Who knows what about me? A survey of behind the scenes personal data sharing to third parties by mobile apps," *Proceeding of Technology Science*, 2015.

[79] A. Zavou, V. Pappas, V. P. Kemerlis, M. Polychronakis, G. Portokalidis, and A. D. Keromytis, "Cloudopsy: An autopsy of data flows in the cloud," in *International Conference on Human Aspects of Information Security, Privacy, and Trust*. Springer, 2013, pp. 366–375.

A The Design Iterations

The designs on the right represent different iterations of designs that were considered during the pilot of the study but were eventually discarded based on feedback from our pilot study. Top to bottom: tabular, matrix, and nodes.

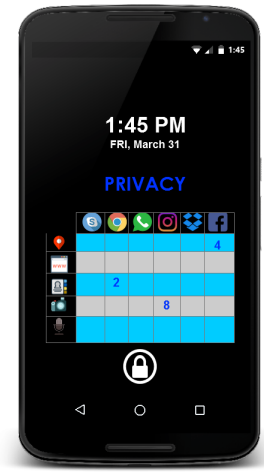

Fig. 3. Tabular design

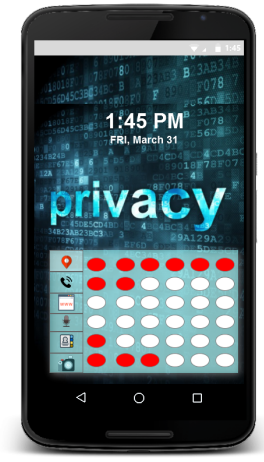

Fig. 4. Matrix design

Fig. 5. Design focused on nodes and links

B Interviews

B.1 Interview protocol

Introduction and consent

"[Good morning/afternoon], my name is ... and I am [title] here at [X] University.

Before we begin, we would like you to review the consent form. Once you've had a chance to read it we may begin when you are comfortable. This document explains your rights as a research participant. During the course of this testing you may be asked some questions that are fairly obvious or repetitive in nature but this is just because I don't want to miss anything. Even if this is the case, I would certainly appreciate your candid responses. If you become uncomfortable, you are free to choose to not answer questions fully or to discontinue the interview. Confidentiality is very important to us. We will not store your name or any identifiable information and we will replace your name with a number. We would like to take an audio recording of this session. Only the investigators of this study will have access to the recordings and it will only be used for transcribing and analytical purposes. Do you give consent to the recording?

[Start recording at this point].

First, thank you for taking the time to speak with me today. I'd like to briefly explain what we are trying to achieve with our study and what you could expect today and if you have any questions please feel free to ask. Our study is exploratory so we will be asking for feedback on a screensaver for a new Android application to look for areas of improvement. Your feedback will be very valuable. In order to help us learn more, I'll show you some designs of the screensaver that we currently have then ask you a few questions. You could expect your entire time here to be about 30-40 minutes."

B.1.1 Overview

Imagine you have just installed a new app on your mobile phone. The app allows you to monitor how the apps installed on your phone (such as Facebook or Dropbox) share your data with other third parties (like advertisers or affiliates). This new app allows you to choose specific apps or specific data types that you would like to focus on and it allows you to see this information as your screensaver. We have a design for this screensaver that we would like your feedback on. First I'll show you the design for 5 seconds then ask for your feedback. After that, I'll show you the design for a longer time and ask for more feedback. Once we are done, I'll ask you to complete an online survey.

B.1.2 First Task (5 seconds): Participants were shown two designs for 5 seconds and implored to remember as much as they could.

(a) Based on the design that you just saw, can you describe what you understand about it?

(b) What do you think is the purpose of the screensaver?

(c) Is there anything that stood out to you?

B.1.3 Second Task (2-3 minutes): Participants were shown the same designs from the first task for 2-3 minutes after which they were implored to remember as much as they could.

(a) Can you describe what you think the screensaver is trying to display?

(b) Is there anything about the design that confuses you?

Wrap up and debrief

(a) Now that you have had the chance to look at both designs, what are your thoughts?
(b) Is there anything else you think I should know?

The participant was thanked for their feedback and time. They were asked to complete a brief survey to collect demographic information and given information about their remuneration.