On Smartphone Users' Difficulty with Understanding Implicit Authentication

Masoud Mehrabi Koushki
University of British Columbia
Vancouver, Canada
mehrabi@ece.ubc.ca

Borke Obada-Obieh
University of British Columbia
Vancouver, Canada
borke@ece.ubc.ca

Jun Ho Huh
Samsung Research
Seoul, Republic of Korea
junho.huh@samsung.com

Konstantin Beznosov
University of British Columbia
Vancouver, Canada
beznosov@ece.ubc.ca

ABSTRACT

Implicit authentication (IA) has recently become a popular approach for providing physical security on smartphones. It relies on behavioral traits (e.g., gait patterns) for user identification, instead of biometric data or knowledge of a PIN. However, it is not yet known whether users can understand the semantics of this technology well enough to use it properly. We bridge this knowledge gap by evaluating how Android's Smart Lock (SL), which is the first widely deployed IA solution on smartphones, is understood by its users. We conducted a qualitative user study (N=26) and an online survey (N=331). The results suggest that users often have difficulty understanding SL semantics, leaving them unable to judge when their phone would be (un)locked. We found that various aspects of SL, such as its capabilities and its authentication factors, are confusing for the users. We also found that depth of smartphone adoption is a significant antecedent of SL comprehension.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **User studies**; *Usability testing*; • **Security and privacy** → **Usability in security and privacy**.

KEYWORDS

Implicit Authentication, Active Authentication, Smartphone Unlocking, Smart Lock, Mental Models

ACM Reference Format:
Masoud Mehrabi Koushki, Borke Obada-Obieh, Jun Ho Huh, and Konstantin Beznosov. 2021. On Smartphone Users' Difficulty with Understanding Implicit Authentication. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3411764.3445386

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05.
https://doi.org/10.1145/3411764.3445386

1 INTRODUCTION

Providing strong physical security for smartphones is becoming more important nowadays. Recent improvements to the capabilities of the phones have allowed their use for a more diverse range of applications, leading users to store highly sensitive data on the phones [16, 47]. This increases the damage (either emotional, financial, or otherwise) that a user would incur in the event of unauthorized access to the device. As such, both manufactures and researchers have ramped up investigation and deployment of stronger solutions for physical security of smartphones [1].

Traditionally, screen locking has been the most prominent of such solutions. Knowledge-based methods—such as passcodes and patterns—were the first methods used to control physical access to mobile devices. However, due to usability [28, 49], memorability [32], and security issues [46, 61, 62] with these methods, manufacturers started to offer biometric unlocking methods (e.g., TouchID and FaceID) as well. These knowledge- and biometrics-based authentication methods are collectively referred to as explicit authentication (EA).

EA methods are still not ideal for controlling access to smartphones. Studies (as recent as 2020) suggest that anywhere between 10% and 35% of smartphone users do not enable any screen locking mechanism on their phones because they find EA inconvenient and cumbersome [16, 30, 45, 48, 53]. Considering that there are nearly 2.5 billion smartphone users around the world [7], this means that at least 200 million smartphones are without security protection.

To address this concern, researchers have proposed implicit authentication (IA) as an alternative to EA. The idea of IA is to identify users through their behavioral traits, such as gait patterns [14], body movements [52], or even eye gaze [34], instead of knowledge-based or biometric input. IA solutions can also use contextual data (such as location [19]) to control access to the device. This eliminates the need for direct user input for every unlocking attempt, resulting in improved usability [36]. Intriguingly, this improvement is shown to be achievable without much sacrifice in security [12, 35–37]. As a result of these promising developments, many IA-based approaches been investigated by academic researchers [14, 50, 52], and IA has seen large-scale commercialization.

Android has been the first major mobile platform to commercialize IA, in the form of Smart Lock (SL). SL can automatically (un)lock devices using contextual signals, such as Bluetooth and GPS, or behavioral traits, such as body movements or gait patterns [22].

As part of the Google Play Services, SL has been deployed on at least 500 million smartphones around the world, making it readily accessible to a considerable portion of smartphone users [23].

Despite the ubiquity of IA, the research community lacks understanding of how well smartphone users understand its semantics. We do not know whether users can predict correctly when IA will (un)lock the phone. Neither do we know whether they understand how IA will work in tandem with EA methods, and to the best of our knowledge, no prior research has investigated these questions. Gaining this insight is important because, as we discuss in Section 5.2, misunderstanding IA can result in unintentionally leaving a smartphone unlocked, undermining the core purpose of IA. Taking as example an IA solution that unlocks the phone at certain "trusted" locations (e.g., home), if the user is unaware of the location detection accuracy, they might leave their phone unsupervised at a semi-public location (e.g., a neighbour's place) presuming that it would be locked, even though it might not be. This can give opportunistic attackers (e.g., social insiders) a chance to snoop on users' data, which is shown to be quite prevalent [47].

In this paper, we bridge this knowledge gap by reporting a study of Smart Lock. We chose SL because not only is it the first widely deployed IA solution on smartphones, but it is also the first to use a combination of EA and IA.[1] This provided us with a unique opportunity to not only study users' understanding of IA, but also to investigate how well the interplay of IA and EA works for users.

We conducted a combination of qualitative user studies, expert reviews, and quantitative surveys. We started by conducting two cognitive walkthroughs of the SL UI, involving 10 HCI-proficient users, to evaluate the learnability of the SL UI. We then conducted 16 think-aloud sessions with average smartphone users (non-experts) to obtain direct user feedback about SL understandability. Finally, we conducted an online MTurk survey (N=331) to validate and refine the findings of our qualitative studies.

Our findings suggest that SL is difficult to understand. For example, we found that the combination of IA with EA makes users confused about which method takes precedence over others and, consequently, when the phone would be (un)locked. This leads to users being uncertain about the state of the phone's physical security. We also discovered that users find it hard to understand which data each IA solution uses for authentication. For example, our participants found it difficult to judge what counts as "motion" for an IA solution that uses this for authentication purposes.

This paper makes the following contributions to the field of IA:

- We provide the first empirical data about users' understanding of IA semantics. We uncover what aspects of SL (a first widely deployed IA solution) can be difficult to understand for the average smartphone user, and how the UI can be designed to minimize confusion.
- We investigate how a combination of IA with EA is understood by smartphone users. We provide insights into the pitfalls of this combination and suggest precautionary measures that should be taken when using it.

[1]For instance, SL can use both implicit location data and explicit facial features of the user to determine if the phone should be unlocked.

2 RELATED WORK

2.1 Smartphone Unlocking

Studies have shown that unlocking increases both the cognitive and the physical overhead of using smartphones, creating inventive for weakening or disabling authentication [31, 56]. It is estimated that smartphones users unlock their devices at least 30 times per day, and each unlocking attempt takes at least 2 seconds to complete [45, 53]. At worst, this time overhead has been shown to consume up to 80% of the whole interaction time with the phone [45]. Harbach et al. [30] estimated that on average, users spend 2.9% of their smartphone interaction time on unlocking [30]. Unlocking is found to be specially cumbersome when using a PIN or password, as there have been reports of usability [49] and memorability [32] issues with these methods.

Considering the overhead, it is not surprising that smartphone users often disable unlocking. It has been reported that between 10% and 35% of smartphone users do not use any unlocking mechanism on their devices, most citing "inconvenience" as the reason [16, 30, 45, 48, 53]. In response to this situation, researchers have proposed to perform implicit authentication (IA) on smartphones.

2.2 IA on Smartphones

The feasibility of smartphone IA is well established. Studies have shown how different modalities (i.e., authentication factors) can successfully be used for this purpose. For example, Frank et al. [18], demonstrated how a classifier could continuously authenticate users based on the way they interact with the touchscreen of a smartphone. Gait patterns [14], body movement [52], biomedical signals [50], and app usage [19] are some other investigated modalities. Some multi-modal IA schemes have also been proposed [64].

There is also substantial empirical evidence for the efficacy of IA. Crawford et al. [12] observed that their participants authenticated 67% less often when using IA, compared to EA. Khan et al. [36] conducted security evaluations and showed several IA schemes to have high identification accuracy and low performance overhead. In terms of usability, Khan et al. [38] reported that 91% of their participants found IA to be convenient.

Despite the promise of IA, which has led users to show interest in adopting this technology [11, 38], there have been no studies conducted on users' understanding of IA semantics. In case of SL, we found studies that investigate the usability of FACE [13] and VOICE [59] (explained below in Section 3). However, these studies do not investigate users' understanding of the semantics of SL.

2.3 Users' Understanding of Security Semantics

It has been well established that correct understanding of computer security semantics is of utmost importance. Firstly, a number of studies have shown how misunderstandings can lead to dangerous security errors. For example, Raja et al. [54] demonstrated that users' incorrect understanding of how the Windows firewall operates can lead to dangerous misconfigurations, potentially exposing users' PCs to remote attacks. Similarly, Chiasson et al. [8] investigated the importance of semantics comprehension when it comes to using password managers, observing that incorrect understandings lead to misconceptions about password security. Secondly, a separate

line of work has demonstrated that incorrect understanding of semantics can hinder effective risk communication. It was shown that security experts and non-experts have different understandings of common security risks [2], and risk communication based on experts' understanding might be ineffective [43].

Users' understanding of smartphone security features has also been studied. For instance, Felt et al. [17] investigated smartphone users' understanding of Android permissions. They found that users have misconceptions about how the permissions work, causing them to be unable to comprehend the risks associated with installing apps. Similarly, Lin et al. [41] observed that users' decisions regarding granting permissions to Android apps depends on their understanding of mobile privacy, which is not uniform. Another example is from the secure messaging domain, where Schroder et al. [57] observed that misunderstanding encryption semantics often prevents the users of SIGNAL [58] from correctly verifying the authenticity of end-to-end encryption keys.

However, as mentioned before, there have been no studies investigating users' comprehension of IA, to the best of our knowledge. For the particular case of Smart Lock, we recently reported an investigation of the spread of its adoption among Android users [48]. We found that only 14% those who owned an SL-capable phone were using SL, and that perceived lack of security was one of the main factors that deterred potential users from SL. We uncovered that users' understanding of SL semantics significantly correlated with this perceived insecurity. However, the authors did not study what aspects of SL could cause confusion for users, or what factors were linked to users' understanding of SL. In this paper, we answer these questions.

This paper uses the same qualitative dataset as Mehrabi et al. [48] (we collected data in the same study sessions with the same participants) and the same methodology for its analysis. However, the focus of analysis for this paper is vastly different than Mehrabi et al., due to differences in research questions (see Section 4). Whereas Mehrabi et al. investigate why users adopt or reject SL, this paper evaluates how users understand SL semantics and what aspects of it can be confusing for them. And while Mehrabi et al. found users' understanding of SL semantics to be correlated with SL adoption, this paper digs deeper and finds how and why those misunderstandings happen. This paper also uses a completely different quantitative dataset than Mehrabi et al. (we conducted separate online surveys with different questions and participants).

3 SMART LOCK OVERVIEW

Smart Lock was first introduced during Google's annual I/O conference keynote in 2014 [21]. In its essence, SL is designed to reduce the number of times users have to unlock their phones, by dismissing the lock screen when the surrounding environment is deemed secure. SL offers the following 5 different methods of unlocking the phone, each of which can be enabled separately:

- **On-body Detection (BODY)** is an IA method that operates based on behavioral traits. It keeps the user's phone unlocked while it is "on-person" (a.k.a., in movement, like running) by detecting the user's body movements and gait patterns. BODY cannot automatically unlock the phone, but will lock it if no movement is detected.
- **Trusted Places (PLACE)** is a contextual[2] IA method that uses GPS signals to automatically unlock the phone at specific locations (e.g., home or work). It will also automatically lock the phone when the device leaves the trusted location.
- **Trusted Devices (DEVICE)** is a contextual IA method that automatically unlocks the user's phone when a designated Bluetooth device is connected to it. It will also automatically lock the phone when the device is disconnected.
- **Trusted Face (FACE)** is an EA method. It allows the user to scan their face to manually unlock the phone. It is not capable of automatically locking or unlocking the phone.
- **Voice Match (VOICE)** is another EA method. It allows the user to say"Ok Google" to manually unlock the phone with their voice. It is not capable of automatically locking or unlocking the phone.

SL is considered an important component of the Android OS, as it is actively advertised on Android-powered smartphones. For example, whenever a new Bluetooth device is paired with such a phone, a notification is shown, encouraging the user to add the device as trusted. Similarly, SL-enabled phones occasionally prompt users to enable BODY or PLACE.

4 STUDY DESIGN

Our research study was designed to answer the following two research questions:

- **RQ1**: How well do Android users understand the semantics of SL? Particularly, what aspects of SL can cause confusion for them?
- **RQ2**: What factors (such as demographics or depth of smartphone adoption) are linked to Android users' understanding of SL?

To answer these questions, we first conducted a qualitative cognitive walkthrough with users. This study, which we describe in detail in Section 5, informed us about how well smartphone users understand SL, and what particular aspects of it might be misunderstood. We used this insight to design our second study, which was an online survey, presented in Section 6, on Amazon Mechanical Turk. The aim of the survey was to verify and expand upon the qualitative findings, leveraging a relatively representative sample of the smartphone user population. All of our data collection and analysis procedures were reviewed and approved by our university's research ethics board.

5 QUALITATIVE STUDY: EXPERT REVIEWS AND THINK-ALOUD SESSIONS

5.1 Methodology

Our first study was a cognitive walkthrough with users (CWU) [24, 42, 44], which we conducted between September 2018 and February 2019. It consisted of two separate parts:

Part I) Cognitive Walkthroughs: To evaluate the learnability of SL (and its UI), we conducted 2 cognitive walkthrough sessions [63] involving 10 HCI-proficient participants who were recruited through the mailing list of our university's HCI research

[2]Meaning it operates based on contextual data rather than behavioral data.

cluster. To qualify as HCI proficient, participants had to have at least 4 months of formal HCI coursework.

In each session, a group of participants (6 in the first session, 4 in the second) walked through the SL UI[3] screen by screen, emulating the actions an SL user would perform to achieve a desired task (e.g., enabling BODY) and discussing any aspects of it could cause confusion. This helped us understand how SL users could potentially perceive its semantics and what potential confusions to look for among the participants of the second part of this study (described below).

To make sure HCI-proficient users would emulate the actions of an SL user realistically, we created a list of common tasks that an SL user might want to do with the UI, with the sequence of actions that they would perform to achieve each task. To compile a comprehensive task list, two researchers used SL on their phones for a period of two weeks. They then collaboratively composed a list of SL UI task affordances and the action sequences for each task. We compared the resulting list with official SL documentation [22], which showed conformance. For presentation to participants, we chose a wording for each task that reflected the goal of the task (e.g., "Enable On-Body Detection"). The full list of tasks is provided in Section 1 of the supplementary materials.

During the sessions, we also asked the participants to give written answers to questions about how each SL method locks or unlocks the phone. These questions were added to the handout (available in Section 1 of the supplementary materials) that we provided to them as part of the cognitive walkthrough protocol. These written answers helped us evaluate whether there were prevalent SL misunderstandings among the HCI-proficient participants.

The average length of the sessions was 130 minutes. Participants were compensated with CAD 20 cash plus refreshments.

We should note that we tested our methodology design before conducting the main 2 cognitive walkthrough sessions described above. We recruited 5 such participants through word of mouth and conducted a 2-hour pilot cognitive walkthrough session. The session showed that our study design was effective, as it identified several confusing aspects of SL that were later confirmed by our main studies. However, we also observed that the length of the session caused fatigue for some participants, reducing their performance. To combat this, we decided to offer refreshments and a break midway through the cognitive walkthrough session. We recruited 2 more HCI-proficient users and conducted a second pilot which showed our modification was effective.

Part II) Think-Alouds: To strengthen our understanding of how smartphone users comprehend SL, we conducted 16 think-aloud sessions with ordinary smartphone users. We did so because, as commonly cited in the literature [44, 63], cognitive walkthroughs lack real user involvement which reduces the ecological validity[4] of their results. We addressed this shortcoming by combining our the cognitive walkthrough findings with the results of think-aloud sessions with ordinary (non-expert) users. This methodology design was inspired by the the Cognitive Walkthrough with Users (CWU) [24] method. We chose this method because think-aloud protocol is found to be a good tool for evaluating comprehension [60, 65].

[3]Section 3 in the supplementary materials of the paper provides screenshots of the main UI screens of each method

[4]Ecological validity refers to the extent to which the experimental conditions of the study are representative of real tasks being done by real users in their natural environment [39]

Table 1: Demographics of the study participants.

Parameter	Property	% (#) of participants	
		CWU (N = 26)	**Survey** (N = 331)
Gender	Female	53.8 (14)	35.3 (117)
	Male	46.2 (12)	64.0 (212)
	Other	0.0 (0)	0.7 (2)
Age	19-24	38.5 (10)	9.7 (32)
	25-34	57.7 (15)	42.6 (141)
	35-44	3.8 (1)	29.6 (98)
	45-54	0.0 (0)	10.0 (33)
	55-64	0.0 (0)	6.9 (23)
	65-74	0.0 (0)	1.2 (4)
Education	< High School	0.0 (0)	0.3 (1)
	High School	7.7 (2)	39.0 (129)
	Bachelor's	38.7 (10)	51.1 (169)
	≥ Master's	53.4 (14)	9.6 (32)
Ethnicity	White	N/A	77.3 (256)
	Black	N/A	8.0 (26)
	Hispanic	N/A	6.0 (20)
	Asian	N/A	6.0 (20)
	Other	N/A	2.7 (9)

In each session, a participant was instructed to perform tasks (the same as the ones we used for cognitive walkthroughs) while thinking out loud about their experience and understanding of SL (we were particularly interested in whether anything confused them about SL). The participants were also given a handout and asked to submit written answers to questions about SL semantics (the handout is provided in Section 1 of the supplementary materials). The sessions were audio recorded and transcribed to accurately capture participants' verbally expressed thoughts. Additionally, while the participant was thinking out loud, two researchers were taking notes about whatever caused confusion.

To test our think-aloud methodology, we recruited 2 participants through word of mouth and conducted 2 pilot think-aloud sessions. While the results showed our approach to be effective in evaluating SL understanding, we occasionally observed "search-and-click" behavior from the participants (they blindly followed instructions without trying to explore and understand SL). To mitigate this issue, we consulted literature on think-aloud protocol [60, 65] and made the following adjustments to the study design:

(1) We would inform participants at the beginning that we would ask them detailed questions about SL semantics. We believed this would motivate them to understand SL.
(2) We would remove the step-by-step task guidance (action sequences) from the handouts and provide participants with the goal of each task only. Note that we retained the step-by-step guidance for cognitive walkthrough (expert) participants to facilitate discussion among experts, as recommend by literature [42, 44].
(3) We would remind participants to think aloud and clarify the reason for their actions, every time they went silent during the sessions.

(4) We would conduct exit interviews with the participants, asking them to verbally clarify their written answers to the SL semantics questions in the handouts, to further gauge their level of SL understanding.

We then conducted two additional pilot sessions. These showed our modifications are effective, as we observed very few instances of "search-and-click."

For the main study, we recruited 16 smartphone users though local Facebook and Craigslist advertising (to be eligible, they had to have owned a smartphone for at least 4 months). The sample size was not fixed from the start. Rather, data collection and analysis were performed concurrently and continued until theoretical saturation was reached (no new codes emerged from the results of the last 2 data collection sessions). The average session length was 40 minutes. We compensated each participant with CAD 30.

The demographics of CWU participants are presented in Table 1. As it shows, our sample was diverse in terms of age, gender, and education. Additionally, we also observed diversity in terms of familiarity with SL.

Data Analysis: The data collected from our CWU study included the transcribed audio recordings of all sessions, participants' written answers to the questions in the handouts, and researchers' field notes. To analyze this data, we used a Thematic Analysis (TA) approach. We chose TA because it is shown to help examine emerging themes from textual data in a transparent and credible way [5, 25, 26]. We followed the steps described in Braun et al. [5]:

First, two researchers coded all the data. To do so, they examined whether a participant's understanding of SL (expressed through either written answers in handouts or verbally expressed thoughts) deviated in some aspect from the ground truth of SL semantics (which we obtained from official SL documentation and verified by our own internal testing). If it did, the researchers coded the text with a label that reflected the aspect of SL that was misunderstood.

Second, the researchers resolved differences in coding through in-person meetings. Differences happened mostly when there were discrepancies between our sources (e.g., between a participant's written and oral answers). In such cases, the researchers present at the CWU session decided which source reflected participants' understanding of SL the best and retained only the codes associated with that source. This process was done iteratively and continued until inter-rater reliability reached a satisfactory 85% (i.e., the coders agreed on which code to use for 85% of the words in the dataset).

Third, each researcher studied all the codes, merging them to draft themes of SL confusion and its antecedents. They then discussed and agreed upon the themes and drafted the results.

5.2 Results

We found SL misunderstandings to be prevalent and specific. Most of the CWU participants had difficulty understanding the semantics of at least some SL methods, as Table 2 shows (additionally, Table 1 in Section 5 of the supplementary materials provides a more detailed view of how participants understood the semantics of each SL method). We observed many confusions not only as they thought out loud when performing tasks, but also when they were asked how they thought each SL method (un)locked the phone.

Misunderstandings, however, were not uniformly distributed across the SL methods. As evidenced by Table 2, we found that PLACE and DEVICE were generally easier to understand, due to most participants having prior experience with Bluetooth and GPS. On the contrary, the unlocking semantics of BODY and the locking semantics of FACE and VOICE were the most confusing. Overall, by thematically analyzing all instances of participants being confused, we identified 4 different categories of SL misunderstanding:

5.2.1 Capabilities of SL. The participants were unsure what SL and its methods were capable of, in terms of locking or unlocking the phone.

Firstly, just the name "Smart Lock" was already confusing. Some participants interpreted it as the ability of the phone to lock itself when the surrounding environment is deemed insecure. However, after interacting with the UI, most of them concluded that SL was rather mostly about unlocking.

The potential for this misinterpretation was first brought up by one of the HCI-proficient participants (code-named P-CW-3) who stated:

> *"... it's called Smart Lock. But it's really more like Smart Unlock because it's not really locking your phone ... it's not as clear about when it actually locks things ... it's more clear about when it unlocks things."*

A think-aloud participant, P-TA-2 (female, 36, software engineer), voiced the same concern by explaining:

> *"... the description [of the SL UI] is really confusing to me ... because it says it keeps your device unlocked when it's safe with you ... [but] I feel like if it's keeping it unlocked, it should be called Smart Unlock, as opposed to Smart Lock. Because when I think Smart Lock, I think it knows when to lock itself. But, the first thing it [SL UI] is talking about is it knows when to keep itself unlocked."*

Interestingly, we observed this misinterpretation to be actually made by some think-aloud participants. P-TA-4 (female, 40, personal trainer) who was non-tech-savvy, believed none of the SL methods could automatically unlock the phone (even after performing all the tasks), citing this naming convention as the reason.

While this was rather a rare example in our data, our findings suggest that overall naming can have implications on users' understanding of IA. The examples above shows that the term "Lock" in "Smart Lock" might give some users the impression that the feature cannot automatically unlock the phone (because they believe it is engineered to lock the phone, not unlock it), which could lead to dangerous errors by users.

Apart from the SL naming, however, we found the discrepancy between the capabilities of SL methods to be even more confusing. The fact that BODY cannot unlock the phone automatically while PLACE and DEVICE can was startling, even to the HCI-proficient participants. Some of them incorrectly believed BODY could both lock and unlock the phone, such P-CW-2 who stated:

> *"when there is someone carrying it, it [BODY] will unlock [the phone] automatically."*

or P-CW-4 who believed:

> *"When I am walking, moving or the phone is in motion [, the phone will unlock]."*

Table 2: Number of CWU participants who correctly understood the un(locking) semantics of each SL method.

Semantics	% (#) participants				
	BODY	**PLACE**	**DEVICE**	**FACE**	**VOICE**
Locking	76.9 (20)	73.1 (19)	57.7 (15)	34.6 (9)	30.8 (8)
Unlocking	30.8 (8)	76.9 (20)	73.1 (19)	88.5 (23)	80.8 (21)

Unsurprisingly, we saw the think-aloud participants make the mistake as well. For example, both P-TA-12 (male, 28, flight attendant) and P-TA-3 (female, 27, unemployed) incorrectly believed that BODY could automatically unlock:

- P-TA-12: *"It [BODY] unlocks the phone when you are in movement, like holding the phone."*
- P-TA-15: *"It [BODY] unlocks the phone while the sensor is on your body and detects motion."*

Therefore, it seems that inconsistency in capabilities is a clear detriment to users' understanding of SL. There might be technological reasons for this inconsistency (e.g., limited accuracy of movement detection). However, as far as the users are concerned, any discrepancy between the capabilities of SL methods causes confusion that could lead to dangerous security errors (e.g., misjudging whether BODY would unlock the phone).

Finally, we found another capabilities-related confusion to be the mixture of IA and EA. Among our participants, this mixture often created unmet expectations about the capabilities of EA methods (FACE and VOICE), which are not capable of automatic un(locking). Evidently, nearly 70% of our think-aloud participants mistakenly believed FACE and VOICE could also automatically lock the phone. For example, when interviewed about FACE semantics, P-TA-1 (male, 31, unemployed) remarked:

> *"[It locks the phone] when someone that's not me looks in the camera."*

Similarly, when asked how VOICE locks the phone, P-TA-9 (female, 22, research assistant) stated:

> *"[It locks the phone] When it doesn't recognize my voice."*

Our further probing with these participants showed that such understandings are caused by VOICE and FACE (which are EA methods) being packaged together with other IA methods in SL. This leads the participants to assume that since DEVICE and PLACE can automatically lock the phone, FACE and VOICE should be capable of it too. Subsequently, participants like P-TA-1 and P-TA-9 try to justify and internalize their understanding by coming up with incorrect explanations like the ones above.

5.2.2 The Modalities (Authentication Factors) of SL. Most participants did not fully understand what kind of data SL used to identify them. As such, they were unsure what they had to do to make SL (un)lock the phone.

In case of BODY, the culprit was the ambiguity of the term "motion." The text presented in the BODY UI describes the feature as being able to keep the phone unlocked for as long as it is in "motion." However, what exactly constitutes "motion" is not clearly communicated. An HCI-proficient participant (P-CW-7) voiced this concern by stating:

> *"... but also 'motion' seems to be the key and I don't understand what kind of motion? ... like when I'm running?"*

Another cognitive walkthrough participant (P-CW-6) similarly stated:

> *"The text [description of 'motion' on [the BODY UI] is not fully clear to me what it means. It says it will be unlocked by the user holding or carrying the device. So, does it mean if I put it in my pocket and [am] moving ... it'll be unlocked?"*

P-CW-6 was specially concerned with this scenario because she believed it could potentially lead to pocket dialing and the related privacy issues.

Another concern regarding "motion" was brought up by P-CW-4. He was wondering whether the "motion" required by BODY needed to be body specific (as the method is named "on-body detection"). He explained:

> *"I'm not sure if it [BODY] will unlock with motion without being on body. It's vague as to whether the phone needs to be on the body or just in motion."*

Overall, we found the concerns with the definition of "motion" valid, as we observed think-aloud participants to exhibit such confusions. As an example, when P-TA-14 (male, 19, student) tried to make the phone lock after setting up BODY, he put it on a nearby desk, but the phone did not lock. He then voiced his frustration by stating:

> *"I expected it [the phone] to [lock] if it's not in my hand ... how would it not lock if it's far away?"*

By probing further, we found that what was unclear to the participant was how BODY detects "motion," what the intensity of the motion should be, and how long the phone takes to detect it. Having this knowledge is important for the participant to make a correct judgement about the state of the phone's security.

Ultimately, as explained by several participants, why "motion" is confusing becomes clearer when we compare BODY to a conventional unlocking method, such as fingerprint. Using fingerprint is fairly straightforward—you put your finger on the sensor and the phone unlocks (and there is no automatic locking). In comparison, it is not very clear how the phone needs to be moved, with what intensity, and for how long, and whether the movement needs to resemble body movement to make BODY function. It's too nuanced. As we saw in the example with P-TA-14, this leads to confusion and frustration.

Lastly, we should note that "motion" was not the only confusing SL modality. DEVICE semantics was also found difficult to fathom.

Some non-tech-savvy think-aloud participants lacked the knowledge to understand how Bluetooth devices are authenticated to each other, concluding that if there is an untrusted device around, the phone will lock. As an example, when we interviewed P-TA-3 (female, 27, unemployed) about when DEVICE would lock the phone, she answered:

> *"... it will lock the phone when you are near a device you have not added as a trusted device."*

When we asked why she thought so, the participant explained that this was simply what she expected from DEVICE, based on what she saw on the UI. She clarified that she had no prior understanding of Bluetooth, and this is what the SL UI led her to believe about DEVICE capabilities. This case showed us that correct understanding of IA methods may sometimes require deeper technical knowledge (e.g., how Bluetooth authentication works) than most ordinary users have.

5.2.3 The Interoperation of SL Methods. How the SL methods interoperate with each other was confusing for the participants. They did not know when the phone would be (un)locked if more than one method was enabled at the same time.

This problem was brought up by several cognitive walkthrough participants, such as P-CW-6 who stated:

> *"Will there be setting conflicts [with SL]? For example, no trusted device [is connected] to my phone, but I'm still at a trusted place; will the phone get locked?"*

Similarly, P-CW-2 remarked:

> *"I [as a user] am kind of confused about how these function[s] work together? Are they independent or overlapped in some ways?"*

This was shown to be a valid concern, as we found SL interoperation to be unclear to most think-aloud participants. P-TA-5 (32, female, health instructor), for example, said the following about PLACE capabilities and its interaction with BODY:

> *"I would guess that when I leave the trusted place, it [the phone] locks. But, I'm not sure how this [PLACE] interacts with on-body detection."*

Interoperation was specially confusing when it was about the interactions between IA and EA methods. As mentioned before, EA methods are not capable of automatically locking the phone. However, when asked when FACE and VOICE lock the phone, some participants thought of SL as a coherent entity and tied the locking capabilities of these EA methods to the IA ones. For example, when asked about how FACE would lock the phone, P-TA-3 stated:

> *"When I am not in motion, the phone would ask me to lock the phone by taking an image of my face."*

Similarly, when we asked P-TA-10 (male, 33, financial consultant) about how VOICE would lock the phone, he answered:

> *"When you set it [the phone] down, as in there is no motion."*

Overall, such observations seem to suggest that interoperation of SL methods is not a matter that is easily understandable by users. Misunderstandings about this can lead to dangerous security errors, such as leaving the house assuming the phone will be locked, whereas it may not be because of another method like DEVICE. Surprisingly, the SL UI does not specifically address how SL methods interoperate at all, leaving it as a guessing game for its users.

5.2.4 The Range Parameters of SL. Understanding when the phone would be (un)locked by PLACE and DEVICE required knowledge of their range parameters, which most of the participants lacked. Our interviews with the think-aloud participants showed that most did not recall the 80-meter operational range of PLACE or the 100-meter one of DEVICE. For example, when we asked P-TA-10 (male, 33, financial consultant) about when DEVICE locks the phone, he stated:

> *"When you take it away from the added trusted device. But how far? Nobody knows!!"*

His exclamation specifically mentioned his lack of knowledge of the range parameter. This knowledge gap may seem insignificant at first. However, it can be essential for correct IA comprehension. This is because these parameters specify the boundaries of security for users. For example, if one user does not know the 80-meter range of PLACE, they might assume that their phone would be locked when not inside their house, where, in actuality, it may not be, when left in their car parked out front.

Interestingly, even if participants knew about the range, PLACE reliability issues sometimes interfered with their correct understanding. It sometimes happened during our study sessions that PLACE failed to function as expected.[5] Such issues caused some participants to doubt their correct understanding of PLACE semantics. For example, when asked how PLACE unlocks the phone, P-TA-14 (male, 19, student) responded:

> *"I don't know. It didn't work and didn't unlock the phone at current location. I expect it to unlock in the room [where study was conducted]."*

Therefore, the data suggests that unreliability is a potential source of confusion. Evidently, even if the semantics of an IA method are clearly conveyed to the user, intermittent operational failures can cause users to doubt their correct understanding, leading to possible dangerous errors (e.g., in case of P-TA-14, the assumption that PLACE cannot automatically unlock the phone).

Regarding range parameters of SL, several of our cognitive walkthrough participants argued that the issue might be the way range parameters are communicated to the user by the SL UI. We discuss this matter further in Section 7.

In the end, to summarize our qualitative findings, our study suggests that SL misunderstandings are common. We found that for users to understand SL correctly, they need to know what each SL method is capable of, what data it uses for authentication, and what its operational parameters are. They also need to know how SL methods interoperate in case more than one is enabled at the same time.

6 QUANTITATIVE STUDY: ONLINE SURVEY

6.1 Methodology

To verify and expand our qualitative findings, we conducted a survey on Amazon Mechanical Turk (MTurk) between August and September 2019. The aim was to leverage a relatively representative

[5] We kept the conditions of the study as similar as possible for all participants. We believe PLACE malfunctions were mainly due to poor GPS signal.

sample of the smartphone user population to evaluate the overall prevalence of SL misunderstandings among smartphone users (especially among SL users, RQ1) and the potential antecedents of these misunderstandings (RQ2). We chose MTurk because it is known to provide quality data for research in usable privacy and security [29, 51, 55].

To recruit participants, we advertised on MTurk, inviting participants with an Android phone to partake in a study about "Smart Lock for Android." We mentioned that participants did not need experience with SL to be eligible for the study (this was because recruiting "SL-novice" participants was necessary for verifying our hypotheses about antecedents of SL understanding, which we discuss later in this section). The survey was only visible to MTurk users who lived in North America[6] and had an approval rating higher than 90%. All those who took the survey were compensated with USD 4.

In the survey, we first asked participants about the following:

(1) **Demographics**: Including age, ethnicity, level of education, and computer background.
(2) **Phone Usage**: How much time they spent using their phones each day, how frequently they unlocked their phones, and the unlocking methods that they had enabled on their devices.
(3) **Privacy-Sensitive App Adoption**: Consisting of 10 Likert-scale questions aimed at evaluating how often participants used privacy-sensitive apps (e.g., social networking) on their phones. The total sum of the scores was used to measure the participants' depth of smartphone adoption. This approach was inspired by Marques et al. [47] (our scale, which is provided Section 2 of the supplementary materials, was a slightly simplified version of theirs).

Next, we provided participants with a quick video introduction to SL. This was done for two reasons: (1) SL is named differently by different phone manufacturers (e.g., it is called Smart Unlock on Huawei phones). The video made it clear to participants what we refer to as SL. (2) To investigate how prior experience with SL correlates with SL understanding (i.e., whether using SL for a period of time helps users understand it better), we intended to contrast SL-experienced participants with SL novices. To this end, we used the video to introduce SL to novices. To avoid inadvertently priming them, however, we carefully crafted the video to limit the amount of information it communicates and make it align with the SL UI.[7]

Afterward, we asked participants whether they knew about SL before participating in our study. And, if so, if they were using any of the SL methods.

Finally, we gauged participants' understanding of SL semantics. We asked them what they thought each SL method was capable of, and how they thought SL methods interoperate (i.e., whether there is an "AND" or "OR" condition for locking or unlocking phones). We did not ask them about range parameters or modalities. This was because our CWU study, as well as our pilot surveys (explained below), showed that this quantitative cross-sectional survey could not sufficiently capture participants' understandings of these aspects, and would not lead to concrete results. The list of all survey questions is provided in Section 2 of the supplementary materials.

To increase the quality of our data, we introduced attention and consistency checks to our survey. These included putting cues in the SL introduction video and asking about them later in the survey (e.g., the video showed participants a word that they needed to input into a text box afterward), checking for inconsistencies in the phone usage answers (e.g., someone claiming to use pattern lock on an iPhone), and checking the response times to see if they deviated significantly from our pilot-based estimate of 15 minutes.

To assess the quality of our survey design, we consulted 7 HCI experts from our university's HCI research cluster. While most aspects of the design were well received by them, some experts were concerned that the video might introduce bias to the results, as the quality of the video content might influence how SL novices understood SL semantics. To address this internal validity risk, we did the following:

Firstly, as mentioned before, we crafted the video solely based on SL UI, so it would not communicate any extra semantic information.

Secondly, we developed an alternative SL introduction medium—a text document[8] augmented with screenshots of the SL UI.

Thirdly, we conducted a pilot study with 10 participants on MTurk using the two introduction mediums (the video and the text document). Results showed that the participants' comprehension of SL semantics was broadly similar to that of CWU participants who experienced the SL UI directly.

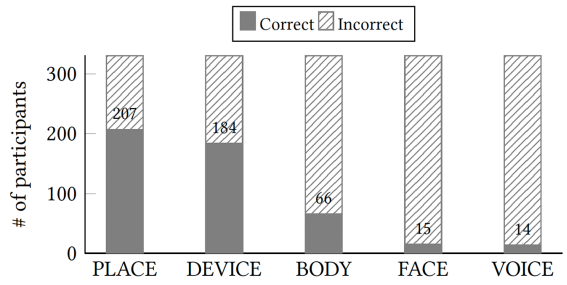
Finally, for the main survey, we randomly assigned survey takers to one of the two mediums and compared SL comprehension among the two groups. A chi-squared test of association revealed no statistically significant differences between the groups ($p > 0.05$).

Overall, the steps we took showed that our SL introduction mediums are fairly representative of the SL UI, in terms of the SL semantics information they communicate. As such, we believe we sufficiently mitigated the risk of the quality and content of the video influencing the outcome of the survey (our consultants agreed).

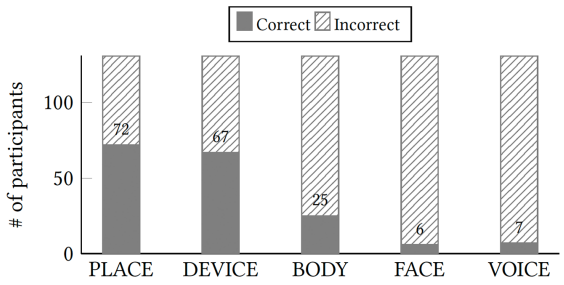
To test the final design, we asked the 10 pilot participants mentioned above to answer all the survey questions and provide written feedback to us. Based on their responses, we made only minor modifications to the wording of some questions.

Afterward, we published the final survey on MTurk and received 382 responses overall. We eliminated 51 answers due to failing the data quality checks mentioned before (10 responses), using IP addresses outside of north America (3), using a VPN or VPS (10), or being flagged as a duplicate or bot by our survey platform Qualtrics (28). The average completion time was 12 minutes. As presented in Table 1, our sample was fairly diverse in age, gender, and education. We also observed diversity in occupation (not included in Table 1 due to great variation). Furthermore, chi-squared tests showed that the distribution of demographics (age, gender, and education) among our participants was not significantly different than that of the US smartphone user population [6]. Hence, we believe our sample to be fairly representative of that population. However, as Table 1 shows, the ethnicity distribution in our sample was

[6]This is a limitation of our study which we will discuss in Section 8.
[7]The video is publicly available on YouTube through this link: https://www.youtube.com/watch?v=N-pC6-kWW0c
[8]The documentation was drafted based on Google's help page for SL. Small modifications were made to make the presentation coherent.



(a) all participants (N = 331)



(b) SL-aware participants (N = 131)

Figure 1: Distribution of survey participants' answers to SL capabilities questions.

Table 3: Results of chi-squared tests of association between having experience with an SL method and answering the corresponding capabilities-related semantics question correctly. V refers to Cramer's V, reflecting the strength of the association.

	BODY	PLACE	DEVICE	FACE	VOICE
Experience	p = 0.227 V = 0.106	p = 0.465 V = 0.063	p = 0.067 V = 0.160	p = 0.626 V = 0.043	p = 0.228 V = 0.105

significantly skewed toward white individuals. This is a limitation of our study that we discuss in Section 8.

Data Analysis: To evaluate the prevalence of SL misunderstandings (and answer RQ1), we used mostly descriptive statistics, which we present in Section 6.2.1. To identify antecedents of SL understanding (RQ2), we used chi-squared tests of association, as we did not assume any particular distribution for the data. The tests were performed for each SL method separately. For each test, correct understanding of the semantics of the SL method (i.e., whether the participant's answer conformed with our established ground truth explained before) was the dichotomous column variable, while each hypothesized antecedent (all of which were categorical) was the row variable. We used Bonferroni correction to mitigate the p-value inflation caused by multiple comparisons. The results are presented in Section 6.2.2.

6.2 Results

6.2.1 Prevalence of SL Misunderstandings. We found SL misunderstandings prevalent, as predicted by our qualitative study. Figure 1a presents how many of the survey participants incorrectly answered the questions about the (un)locking capabilities of each SL method. About 80% of the participants incorrectly answered the questions about the capabilities of FACE, VOICE, and BODY. Based on our CWU findings, we predicted that the majority would be confused about these methods, due to the IA-EA mixture. The survey clearly confirmed this prediction.

Interestingly, our data also suggests that SL-aware participants (those who reported having known about SL before participating in our study) were not more likely to answer the questions correctly. As Figure 1 shows, the distribution of correct vs incorrect answers were nearly identical between the SL-aware subgroup and the overall sample. Furthermore, as presented in Table 3, chi-squared tests showed no statistically significant (p-values > 0.05) correlation between a participant reporting experience with an SL method and answering the corresponding semantics question correctly. This was an extension to our CWU results, as we did not have enough SL-experienced participants in our qualitative study to make such observations. This new finding suggests that SL misunderstandings are prevalent even among experienced users.

For BODY, FACE, and VOICE (where most participants answered incorrectly), the common mistake among participants was believing that these methods could unlock the phone automatically. This is in line with our CWU findings. As we discussed in Section 5.2, the confusion seems to be the mixture of IA and EA creating unmet expectations about FACE and VOICE capabilities, and the discrepancy between the capabilities of BODY and other IA methods.

As for PLACE and DEVICE, we see in Figure 1 that the majority (nearly 60%) answered the semantics questions correctly. This was again in line with the CWU findings (see Table 2), as those methods were found easier to understand. However, a new observation was that the prominent mistake among the minority was believing PLACE and DEVICE could not lock the phone automatically. Based on our CWU results, we conjecture this to be due to the participants extending their understanding of traditional unlocking methods (e.g., PIN or fingerprint) to IA, believing SL could not lock the phone. However, further studies are needed to firmly confirm this conjecture.

Finally, for the interoperation of the SL methods, we found that only 9.7% (N=32) of participants answered the corresponding semantics question incorrectly. The other 74% (N=245) correctly assumed an OR logical relationship, even though it was never communicated to them. This observation seems to suggest that OR is what the participants expect by default, which is, evidently, what SL provides. However, even though the survey data shows that interoperation misunderstandings are not highly prevalent, we still

Table 4: Results of chi-squared tests of association between SL understanding and our anticipated antecedents of it. Significant p-values are underscored (assuming $\alpha = 0.05$). V refers to Cramer's V, reflecting the strength of the association.

	BODY	PLACE	DEVICE	FACE	VOICE
Age	p = 0.756 V = 0.076	p = 0.103 V = 0.153	p = 0.705 V = 0.081	p = 0.485 V = 0.102	p = 0.527 V = 0.099
Computer Literacy	p = 0.245 V = 0.064	p = 0.390 V = 0.047	p = 0.133 V = 0.082	p = 0.642 V = 0.026	p = 0.724 V = 0.019
Privacy App Adoption	p = 0.380 V = 0.076	p = 0.017 V = 0.156	p = 0.035 V = 0.142	p = 0.356 V = 0.079	p = 0.066 V = 0.128
Security Proficiency	p = 0.122 V = 0.113	p = 0.250 V = 0.091	p = 0.379 V = 0.076	p = 0.494 V = 0.065	p = 0.171 V = 0.103

believe that the corresponding semantics should be clearly communicated by the UI. As our CWU findings show, misunderstandings can lead to various dangerous errors by the users, be it among only 10% of the participants (conservatively).

To answer **RQ1** then, our survey results suggest that misunderstandings about SL are prevalent. They are mostly regarding the capabilities of SL methods, with occasional confusions about SL interoperation as well.

6.2.2 Antecedents of SL Understanding. Per **RQ2**, we were interested to see if pre-existing factors (e.g., demographics or depth of smartphone adoption) were linked to correct SL understanding. We collected a multitude of such data in our survey. However, to avoid a fishing expedition, we only examined the effect of factors that were referred to by the related work.

We should note that our list of antecedents is not comprehensive, not that we intended for it to be. We only studied how our research aligns with some notable related work from authors who investigated the antecedents of the understanding of conventional unlocking methods. While our study provides a first insight into this topic, future studies are needed to study the matter further.

Furthermore, we should also note that since we did not observe any significant differences in SL comprehension between novice and experience users, we included all participants in the following analysis. Had we found any differences, we would have included SL-aware participants only.

To start with, we examined whether age was a factor that was linked to SL understanding. Age is shown to be associated with smartphone users' perception of conventional phone unlocking methods [53]. As such, we were interested to see if it had a similar correlation with SL understanding, as well. On the contrary, our data analysis revealed no significant association between age and correct understanding of SL capabilities (See Table 4) or SL Interoperation (p = 0.595, Cramer's V = 0.092). This observation suggests that correct SL comprehension is not dependent solely on cognitive capability which could potentially give younger individuals an advantage.

Next, we examined the link between depth of smartphone adoption and SL comprehension. Previously, Marques et al. [47] showed that power phone users (those who use their phones for a more diverse range of applications) usually have better understandings of phone security features. We were interested to see if this was the case with SL, as well. Our analysis showed that it is, in fact, so. We observed statistically significant associations between privacy app adoption score and correct understanding of PLACE and DEVICE capabilities (see Table 4). However, no similar association was detected for BODY, FACE, or VOICE capabilities (Table 4) or SL interoperation (p = 0.339, V = 0.081). This was in line with our previous findings. Evidently, correct comprehension of PLACE and DEVICE requires deeper Bluetooth and GPS knowledge, which power phone users tend to possess.

Thirdly, we investigated the link between computer literacy and SL understanding. In particular, we examined whether those who had a computer-related occupation were less likely to exhibit the confusions we discussed in the previous section. We did this examination because such correlations have been observed by other studies [2, 41]. Our results, however, did not suggest any such association (see Table 4). As such, in confirmation of our CWU findings, it seems that SL semantics require specific types of knowledge that goes beyond typical computer literacy.

Lastly, we should note that we did not observe any association between medium of introduction to SL, and SL comprehension. As mentioned in Section 6.1, we added a second SL introduction medium (a textual presentation) to our study, to make sure the video does not bias the results. To this end, one half of the participants saw the video, while the other read the document. Chi-squared test revealed no association between the medium of introduction and correct answers to any of the SL semantics questions (p = 0.735, V = 0.018). As such, the video seemed to had not caused any significant bias in the data.

In summary then, to answer **RQ2**, our survey data suggests that, in agreement with the related work, depth of smartphone adoption is a significant antecedent of SL understanding. However, contrary to the related work, we did not find any association between age or computer literacy with SL comprehension.

7 DISCUSSION

Research suggests that IA is a promising technology for providing better physical security protection on smartphones. Not only can it make unlocking more convenient [12, 38], but it does so with only minor sacrifices to security [36]. Even more, users are actually willing to adopt this new technology, provided that it is implemented well [11, 38, 48].

However, it is becoming clear that hitting the sweet spot could be very tricky. Not only must IA schemes address numerous security challenges (e.g., resistance to mimicry attacks and minimal authentication delay) [36], but they also need to resolve the issue of intermittently available data (i.e., when to switch to explicit authentication) [38], which can adversely affect the user experience. To add to these challenges, this paper discovers a new challenge for IA to overcome, which is to efficiently communicate its semantics to users. Our SL case study vividly demonstrates how important this issue is for a successful IA deployment, yet how difficult it is to address.

We found that complex SL semantics confuse users. There are a multitude of aspects to SL that can cause misunderstandings and dangerous security errors. Firstly, the users might misunderstand what type of data SL uses to authenticate them (e.g., how is motion detected by BODY?, see Section 5.2.2). Secondly, they might not know what each SL method is capable of (e.g., can BODY automatically unlock the phone?, see Section 5.2.1). Thirdly, they might not be aware of how SL methods work together (e.g., what happens if both PLACE and BODY are enabled?, see Section 5.2.3). And, lastly, users probably won't have knowledge of the parameters of the SL methods (e.g., how far should the user be from their home, for PLACE to lock their phone?, see Section 5.2.4). Any of these misunderstandings can lead users to misjudge when their phone would be (un)locked, potentially leading to unauthorized access to their sensitive data.

These misunderstandings are more than hypothetical. They are prevalent. Nearly 80% of our 331 survey participants overestimated the capabilities of some SL methods, namely BODY, FACE, and VOICE (see Section 6.2.1). They incorrectly believed that these methods could automatically unlock the phone. Similarly, as also discussed in Section 6.2.1, a considerable number of the participants (40%) incorrectly believed that PLACE and DEVICE could not lock the phone automatically, expecting these unlocking methods to behave like traditional ones (e.g., PIN or fingerprint). And lastly, several participants (nearly 10%) misunderstood how SL methods would interoperate, since there is no explicit indication of this in SL UI.

SL misunderstandings are also universal. We could not identify any demographic group of users who would understood SL better (see Section 6.2.2). We found that younger and older adults misunderstood SL alike. We also found that computer literacy did not necessarily translate to better SL comprehension. The only factor that we found correlating with SL understanding was depth of smartphone adoption, which suggests that only those who have specific knowledge are more likely to understand SL correctly.

Our findings suggest that the SL UI (see screenshots in Section 3 and 4 of the supplementary materials) is partially responsible for these confusions. It is vague about the capabilities of each method (see Section 5.2.1), it does not clearly define what "motion" entails in the context of BODY (see Section 5.2.2), and it is not clear about why it is communicating the range parameters of DEVICE and PLACE (see Section 5.2.4). Worse, it does not communicate at all how SL methods interoperate (see Section 5.2.3). These deficiencies require participants to manually explore and experiment with SL, in order to gauge the validity of their initial impressions, which our survey showed to be error-prone (as discussed in Section 6.2.1, having experience with an SL method made no difference in understanding its semantics correctly). To mitigate these misunderstandings, we rely on our findings to offer the following recommendations for IA UX design on smartphones:

1) COMMUNICATE CLEARLY WHAT DATA EACH IA METHOD USES AND HOW: We recommend that the type of data each IA method uses for authentication purposes should be clearly communicated to users. And, in this regard, no deep technical knowledge should be assumed on the users' part.

Taking BODY as an example, we believe how it detects motion, how big the movement should be, and how long should this movement last are essential to convey in some way, even though these specificities might not seem important at first. For example, our results (Section 5.2.1) suggest that the lack of this understanding can result in users leaving their phone unprotected because they might not know whether being in a moving car counts as motion or not.

Another example is DEVICE, where we believe that the UI should specifically address how it would identify the trusted device (e.g., based on some hardware serial number). One might think that this information is common knowledge and need no explicit explanation. However, we found (in Section 5.2.2) this to be not the case, as some participants lacked this knowledge.

While it is out of scope of this paper to explore how to effectively communicate this information, the way the SL UI does it seems to be sub-optimal, to say the least. In the case of BODY, for example, the main UI lacks any of the information that we suggest to communicate to the user. A help page accessible through an obscure button (that only 2 of our CWU participants clicked on) is the only place that explains how BODY works.

One efficient way of conveying the semantics could be animations, as they have been shown to be effective in communicating meaning [4], specifically in information security [3]. Animations were also suggested by some of our CWU participants (both, HCI-proficient and ordinary smartphone users), e.g.,

> *"I think for me, it would be better to have animations to show how it [BODY] works ... this graphic [still image on BD UI] doesn't tell me much as to how Smart Lock on-body detection works."* [P-TA-9]

2) CLEARLY STATE THE CAPABILITIES OF IA: We recommend that the unlocking capabilities of each IA method (i.e., whether it can automatically lock or unlock the phone) should be clearly addressed and communicated to the user by the UI.

As we saw with the case of SL, this is particularly important when IA and EA methods are mixed together (see Section 5.2.1 and 6.2.1). Evidently, most participants incorrectly believed that FACE or VOICE are capable of automatically locking the phone. As discussed in Section 5.2.3, we believe that the interplay of FACE and VOICE (EA methods) with BODY, PLACE, and DEVICE (IA methods) resulted in this incorrect understanding of EA capabilities.

The current SL UI does not communicate the capabilities of each method explicitly. The user has to deduce it from the descriptions in the UI (see Screenshots in Section 3 of supplementary materials). Fortunately, it is rather easy to communicate whether an IA method

is capable of locking or unlocking or both. Exactly how this information can be communicated effectively is of course out of scope of this paper. However, a suggestion by some of our HCI-proficient participants was to name the IA methods in a manner that clearly distinguishes them from the EA ones. P-CW-2, for example, stated:

> *"I think that the names like Trusted Face and Trusted devices are similar, [even though] they function differently. So maybe you can change the name of one?"*

3) MAKE INTEROPERATION SEMANTICS CLEAR: We recommend that the UI communicate clearly how each IA scheme interoperates with other IA or EA schemes on the phone (i.e., whether this is an AND or OR logical relationship between them). Otherwise, users might mistakenly assume that one authentication method takes precedence over the other, resulting in dangerous errors.

The results of our survey study (see Section 6.2.1) showed that most users expect an OR condition by default, which is evidently what SL provides. However, the SL UI never communicates this matter explicitly, leading some users (at least 10% according to our findings in Section 6.2.1) to incorrectly assume an AND relationship, which can lead to security errors.

Additionally, the fact that 10% of survey participants expected an AND relationship suggests that the matter of how SL methods should interoperate may be subjective, and therefore should be left to the user as a configuration option. This was also suggested by some of our HCI-proficient participants. In either case, our data clearly shows that the UI cannot remain silent on this matter and should address it to avoid dangerous errors.

4) PROVIDE ACCURATE AND RELIABLE VALUES FOR PARAMETERS: The operational parameters of IA methods (e.g., the range of Bluetooth or GPS connections) should be clearly communicated by the UI. As discussed in Section 5.2, knowledge of these parameters is important for users to make sound security judgements.

However, our CWU study suggests that mere presence of these parameters in the UI is not an optimal way of communicating them. In the SL UI, this information is presented through several screens and warning messages (see screenshots in Sections 3 and 4 of the supplementary materials), which most of our participants were either confused by or dismissed as unimportant.

As a case in point, the range parameter in DEVICE UI is communicated through a note (see Figure 1d in the supplementary materials) without any context as to why this information is presented. We believe that this is why our participants often dismissed it as unimportant (Section 5.2).

Our findings in this regard seem to align somewhat with the case of privacy policy communications, which have been studied extensively. It has been shown that most users ignore text-based policy statements because they find them irrelevant or difficult to understand [10, 15]. Subsequently, several approaches have been proposed to improve privacy policy statements. The most notable [27, 40] provides context (e.g., how the data will be used) and appears to be somewhat successful [15]. Interestingly, we saw in our CWU study that the example use cases provided by DEVICE UI (e.g., that you can use it with your car's Bluetooth system) were well received by our participants. Therefore, providing example use cases might be a promising way to communicate range parameters.

And, lastly, we should note that, as discussed in Section 5.2, unreliable operation (e.g., inconsistent range of GPS) can cause users to doubt their understanding of the semantics, even when it is correct. Therefore, the UI needs to be wary of this fact as well and provide the necessary warnings to the user.

8 LIMITATIONS

Any generalizations of our findings need to be performed carefully due to the following study limitations:

Firstly, the cross-sectional design of our CWU study prevented us from fully investigating the effect of prolonged SL usage on participants' understating of its semantics. It is possible that continued exposure to SL can cause certain misconceptions that our CWU study was unable to capture. Our survey data bridged this gap somewhat (we surveyed long-term SL adopters). However, more longitudinal studies (e.g., diary studies) are needed to further investigate this matter.

Secondly, similar to other studies on smartphone usage [9, 13, 30, 33], our survey sample was not fully representative of the smartphone user population. It is shown that cultural factors affect smartphone users' unlocking behavior and attitude [29]. However, due to limited resources, we only included North American (NA) participants in our studies. Therefore, our results are only generalizable to NA smartphone users. Similarly, the ethnicity distribution among our participants was heavily skewed toward whites, which is a common limitation of MTurk studies [51, 55], especially the ones on smartphones [9, 13, 30, 33]. This limits the generalizability of our results. However, we still believe that our findings are valuable, as a first insight into smartphone users' understanding of IA.

Finally, our participants self-reported their prior experiences with SL. Thus, as is common with self-reported data, it is possible that the participants' answers were not completely reflective of their reality [20]. While we eliminated those responses that showed clear inconsistencies, it is still possible that self-reporting has affected the quality or our data, and hence the results.

9 CONCLUSION

While implicit authentication (IA) is becoming a popular approach for protecting physical security of smartphones, not much research has been done in investigating/improving users' understanding of this technology. To bridge this knowledge gap, we used Smart Lock (the first widely deployed IA scheme) as a case study and evaluated how its semantics are understood by Android users. We found SL misunderstandings to be prevalent, universal, and mostly due to insufficient communications by the SL UI. We identified four aspects of SL that have the most potential for being confusing, namely the authentication modalities, the capabilities of IA, interoperation of IA and EA, and the operating parameters of IA. Based on the findings, we provided a set of recommendations on how to improve users' understanding of these aspects. Ultimately, our goal is to facilitate wider deployment of IA on smartphones, as it is known to provide a better user authentication experience compared to traditional phone unlocking methods.

ACKNOWLEDGMENTS

This research has been supported in part by a research grant from Samsung and by a gift from Scotiabank to the University of British Columbia. We would like to thank members of the Laboratory for Education and Research in Secure Systems Engineering, who provided their feedback on the reported research and earlier versions of the paper. We thank our anonymous reviewers for all the feedback and suggestions they provided to improve the paper. Stylistic and copy editing by Eva van Emden helped to improve readability of this paper.

REFERENCES

[1] Nasser O Alshammari, Alexios Mylonas, Mohamed Sedky, Justin Champion, and Carolin Bauer. 2015. Exploring the adoption of physical security controls in smartphones. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*. Springer, Cham, Lisboa, Portugal, 287–298.

[2] Farzaneh Asgharpour, Debin Liu, and L Jean Camp. 2007. Mental Models of Computer Security Risks.. In *Workshop on the Economics of Information Security*. WEIS, Pittsburgh, USA, 1–9.

[3] Cheuk Hang Au, Kyle CS Lam, Walter SL Fung, and Xin Xu. 2016. Using animation to develop a MOOC on Information Security. In *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE, Bali, Indonesia., 365–369.

[4] Ronald Baecker, Ian Small, and Richard Mander. 1995. Bringing icons to life. In *Readings in Human–Computer Interaction*. Elsevier, San Francisco, CA, USA, 444–449.

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[6] Pew Research Center. 2019. Mobile Technology and Home Broadband 2019. https://www.pewinternet.org/2019/06/13/mobile-technology-and-home-broadband-2019/. Accessed: 2019-07-26.

[7] Pew Research Center. 2019. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/. Accessed: 2020-09-14.

[8] Sonia Chiasson, Paul C van Oorschot, and Robert Biddle. 2006. A Usability Study and Critique of Two Password Managers.. In *USENIX Security Symposium*, Vol. 15. USENIX Association, Vancouver, BC, Canada, 1–16.

[9] Geumhwan Cho, Jun Ho Huh, Soolin Kim, Junsung Cho, Heesung Park, Yenah Lee, Konstantin Beznosov, and Hyoungshick Kim. 2020. On the Security and Usability Implications of Providing Multiple Authentication Choices on Smartphones: The More, the Better? *ACM Transactions on Privacy and Security (TOPS)* 23, 4 (2020), 1–32.

[10] Lorrie Faith Cranor. 2012. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *J. on Telecomm. & High Tech. L.* 10 (2012), 273.

[11] Heather Crawford and Karen Renaud. 2014. Understanding user perceptions of transparent authentication on a mobile device. *Journal of Trust Management* 1, 1 (2014), 7.

[12] Heather Crawford, Karen Renaud, and Tim Storer. 2013. A framework for continuous, transparent mobile device authentication. *Computers & Security* 39 (2013), 127–136.

[13] Alexander De Luca, Alina Hang, Emanuel Von Zezschwitz, and Heinrich Hussmann. 2015. I feel like I'm taking selfies all day!: towards understanding biometric authentication on smartphones. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul, Korea, 1411–1414.

[14] Mohammad Omar Derawi, Claudia Nickel, Patrick Bours, and Christoph Busch. 2010. Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. IEEE Computer Society, Washington, DC, USA, 306–311.

[15] Nico Ebert, Kurt Alexander Ackermann, and Peter Heinrich. 2020. Does Context in Privacy Communication Really Matter?—A Survey on Consumer Concerns and Preferences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Virtual Conference, 1–11.

[16] Serge Egelman, Sakshi Jain, Rebecca S Portnoff, Kerwell Liao, Sunny Consolvo, and David Wagner. 2014. Are you ready to lock?. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, Scottsdale, Arizona, USA, 750–761.

[17] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the eighth symposium on usable privacy and security*. ACM, Washington, DC, USA, 1–14.

[18] Mario Frank, Ralf Biedert, Eugene Ma, Ivan Martinovic, and Dawn Song. 2013. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE transactions on information forensics and security* 8, 1 (2013), 136–148.

[19] Lex Fridman, Steven Weber, Rachel Greenstadt, and Moshe Kam. 2017. Active authentication on mobile devices via stylometry, application usage, web browsing, and GPS location. *IEEE Systems Journal* 11, 2 (2017), 513–521.

[20] Robert M Gonyea. 2005. Self-reported data in institutional research: Review and recommendations. *New directions for institutional research* 2005, 127 (2005), 73–89.

[21] Google. 2019. Google I/O 2014 Keynote. https://www.youtube.com/watch?time_continue=1659&v=biSpvXBGpE0. Accessed: 2019-02-14.

[22] Google. 2019. Use Google Smart Lock. https://bit.ly/2XTnGeG. Accessed: 2020-06-12.

[23] Google Play Services. 2020. Number of Installs. https://play.google.com/store/apps/details?id=com.google.android.gms&hl=en_CA. Accessed: 2020-05-01.

[24] T Granollers and J Lorés. 2005. Cognitive Walkthrough With Users: an alternative dimension for usability methods. In *Proc. HCI International*. CRC Press, Taylor & Francis Group, Las Vegas, Nevada, USA, 1–8.

[25] Thomas RG Green, Margaret M Burnett, Andrew Jensen Ko, Karen J Rothermel, Curtis R Cook, and Justin Schonfeld. 2000. Using the cognitive walkthrough to improve the design of a visual programming experiment. In *Visual Languages, 2000. Proceedings. 2000 IEEE International Symposium on*. IEEE, Washington, DC, USA, 172–179.

[26] Greg Guest, Kathleen M MacQueen, and Emily E Namey. 2011. *Applied thematic analysis*. sage, CA, USA.

[27] Margaret Hagen. 2016. User-centered privacy communication design. In *Twelfth Symposium on Usable Privacy and Security ({SOUPS}*. USENIX Association, Denver, CO, USA, 1–7.

[28] Marian Harbach, Alexander De Luca, and Serge Egelman. 2016. The anatomy of smartphone unlocking: A field study of android lock screens. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose CA, USA, 4806–4817.

[29] Marian Harbach, Alexander De Luca, Nathan Malkin, and Serge Egelman. 2016. Keep on lockin'in the free world: a multi-national comparison of smartphone locking. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose, CA, USA, 4823–4827.

[30] Marian Harbach, Emanuel Von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 2014. It's a hard lock life: A field study of smartphone (un) locking behavior and risk perception. In *Symposium on usable privacy and security (SOUPS)*. USENIX & ACM SIGCHI, Menlo Park, CA, USA, 213–230.

[31] Eiji Hayashi, Oriana Riva, Karin Strauss, AJ Brush, and Stuart Schechter. 2012. Goldilocks and the two mobile devices: going beyond all-or-nothing access to a device's applications. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*. ACM, Washington, DC, USA, 2.

[32] Jun Ho Huh, Hyoungshick Kim, Rakesh B Bobba, Masooda N Bashir, and Konstantin Beznosov. 2015. On the memorability of system-generated PINs: Can chunking help?. In *Eleventh Symposium On Usable Privacy and Security ({SOUPS}*. USENIX Association, Ottawa, ON, Canada, 197–209.

[33] Jun Ho Huh, Saurabh Verma, Swathi Sri V Rayala, Rakesh B Bobba, Konstantin Beznosov, and Hyoungshick Kim. 2017. I Don't Use Apple Pay because it's less secure...: perception of security and usability in mobile tap-and-pay. In *Workshop on Usable Security*, Vol. 12. Internet Society, San Diego, CA, 1–12.

[34] Christina Katsini, Yasmeen Abdrabou, George E Raptis, Mohamed Khamis, and Florian Alt. 2020. The Role of Eye Gaze in Security and Privacy Applications: Survey and Future HCI Research Directions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Virtual Conference, 1–21.

[35] Hassan Khan. 2016. *Evaluating the Efficacy of Implicit Authentication Under Realistic Operating Scenarios*. Ph.D. Dissertation. University of Waterloo.

[36] Hassan Khan, Aaron Atwater, and Urs Hengartner. 2014. A comparative evaluation of implicit authentication schemes. In *International Workshop on Recent Advances in Intrusion Detection*. Springer, Gothenburg, Sweden, 255–275.

[37] Hassan Khan, Aaron Atwater, and Urs Hengartner. 2014. Itus: an implicit authentication framework for android. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, Maui Hawaii, USA, 507–518.

[38] Hassan Khan, Urs Hengartner, and Daniel Vogel. 2015. Usability and Security Perceptions of Implicit Authentication: Convenient, Secure, Sometimes Annoying.. In *Eleventh Symposium On Usable Privacy and Security ({SOUPS}*. USENIX Association, Ottawa, ON, Canada, 225–239.

[39] Suzanne Kieffer, Ugo Braga Sangiorgi, and Jean Vanderdonckt. 2015. Ecoval: A framework for increasing the ecological validity in usability testing. In *Forty-eighth Hawaii International Conference on System Sciences*. IEEE, Hawaii, USA, 452–461.

[40] Alfred Kobsa and Maximilian Teltzrow. 2004. Contextualized communication of privacy practices and personalization benefits: Impacts on users' data sharing and purchase behavior. In *International Workshop on Privacy Enhancing Technologies*. Springer, Toronto, ON, Canada, 329–343.

[41] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, Pittsburgh, Pennsylvania, USA, 501–510.

[42] Wallace Lira, Renato Ferreira, Cleidson de Souza, and Schubert Carvalho. 2014. Experimenting on the cognitive walkthrough with users. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. ACM, Toronto, ON, Canada, 613–618.

[43] Debin Liu, Farzaneh Asgharpour, and L Jean Camp. 2008. Risk communication in security using mental models. *Usable Security* 7 (2008), 1–12.

[44] Thomas Mahatody, Mouldi Sagar, and Christophe Kolski. 2010. State of the art on the cognitive walkthrough method, its variants and evolutions. *Intl. Journal of Human–Computer Interaction* 26, 8 (2010), 741–785.

[45] Ahmed Mahfouz, Ildar Muslukhov, and Konstantin Beznosov. 2016. Android users in the wild: Their authentication and usage behavior. *Pervasive and Mobile Computing* 32 (2016), 50–61.

[46] Philipp Markert, Daniel V Bailey, Maximilian Golla, Markus Dürmuth, and Adam J AviG. 2020. This PIN Can Be Easily Guessed: Analyzing the Security of Smartphone Unlock PINs. In *IEEE Symposium on Security and Privacy (SP)*. IEEE, Virtual Conference, 286–303.

[47] Diogo Marques, Ildar Muslukhov, Tiago Guerreiro, Luís Carriço, and Konstantin Beznosov. 2016. Snooping on mobile phones: Prevalence and trends. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, Vol. 2. USENIX Association, Denver, CO, USA, 77.

[48] Masoud Mehrabi Koushki, Borke Obada-Obieh, Jun Ho Huh, and Konstantin Beznosov. 2020. Is Implicit Authentication on Smartphones Really Popular? On Android Users' Perception of "Smart Lock for Android". In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Oldenburg, Germany, 1–17.

[49] William Melicher, Darya Kurilova, Sean M Segreti, Pranshu Kalvani, Richard Shay, Blase Ur, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Michelle L Mazurek. 2016. Usability and security of text passwords on mobile devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose, CA, USA, 527–539.

[50] Arsalan Mosenia, Susmita Sur-Kolay, Anand Raghunathan, and Niraj K Jha. 2017. CABA: Continuous authentication based on BioAura. *IEEE Trans. Comput.* 66, 5 (2017), 759–772.

[51] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70 (2017), 153–163.

[52] Abena Primo, Vir V Phoha, Rajesh Kumar, and Abdul Serwadda. 2014. Context-aware active authentication using smartphone accelerometer measurements. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. IEEE, San Juan, Puerto Rico, USA, 98–105.

[53] Lina Qiu, Alexander De Luca, Ildar Muslukhov, and Konstantin Beznosov. 2019. Towards Understanding the Link Between Age and Smartphone Authentication. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow, UK, 1–10.

[54] Fahimeh Raja, Kirstie Hawkey, and Konstantin Beznosov. 2009. Revealing hidden context: improving mental models of personal firewall users. In *Proceedings of the 5th Symposium on Usable Privacy and Security*. ACM, Mountain View, CA, USA, 1.

[55] Elissa M Redmiles, Sean Kross, and Michelle L Mazurek. 2019. How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples. In *IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 1326–1343.

[56] Oriana Riva, Chuan Qin, Karin Strauss, and Dimitrios Lymberopoulos. 2012. Progressive Authentication: Deciding When to Authenticate on Mobile Phones.. In *USENIX Security Symposium*. USENIX Association, Bellevue, WA, USA, 301–316.

[57] Svenja Schröder, Markus Huber, David Wind, and Christoph Rottermanner. 2016. When SIGNAL hits the fan: On the usability and security of state-of-the-art secure mobile messaging. In *European Workshop on Usable Security. IEEE*. IEEE, Saarbrücken, Germany, 1–7.

[58] Signal. 2020. Signal App. https://signal.org/en/. Accessed: 2020-06-12.

[59] Shari Trewin, Cal Swart, Larry Koved, Jacquelyn Martino, Kapil Singh, and Shay Ben-David. 2012. Biometric authentication on a mobile device: a study of user effort, error and task disruption. In *Proceedings of the 28th Annual Computer Security Applications Conference*. ACM, Orlando, Florida, USA, 159–168.

[60] Suzanne E Wade. 1990. Using think alouds to assess comprehension. *The Reading Teacher* 43, 7 (1990), 442–451.

[61] Ding Wang, Qianchen Gu, Xinyi Huang, and Ping Wang. 2017. Understanding human-chosen pins: characteristics, distribution and security. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*. ACM, Abu Dhabi, United Arab Emirates, 372–385.

[62] Ding Wang, Zijian Zhang, Ping Wang, Jeff Yan, and Xinyi Huang. 2016. Targeted online password guessing: An underestimated threat. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. ACM, Vienna, Austria, 1242–1254.

[63] Chauncey Wilson. 2013. *User interface inspection methods: a user-centered design method*. Newnes, Waltham, MA, USA.

[64] Heiko Witte, Christian Rathgeb, and Christoph Busch. 2013. Context-aware mobile biometric authentication based on support vector machines. In *2013 Fourth International Conference on Emerging Security Technologies*. IEEE, Washington, DC, USA, 29–32.

[65] Kirsty A Young. 2005. Direct from the source: the value of'think-aloud'data in understanding learning. *Journal of Educational Enquiry* 6, 1 (2005), 19–33.