Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Erratum or Corrigendum

"It's all about the start" classifying eyes-free mobile authentication techniques

Flynn Wolf[a], Adam J. Aviv[b,*], Ravi Kuber[a]

[a] *University of Maryland, Baltimore County, United States*
[b] *United States Naval Academy, United States*

ABSTRACT

Mobile device users avoiding observational attacks and coping with situational impairments may employ techniques for eyes-free mobile unlock authentication, where a user enters his/her passcode without looking at the device. This study supplies an initial description of user accuracy in performing this authentication behavior with PIN and pattern passcodes, with varying lengths and visual characteristics. Additionally, we inquire if tactile-only feedback can provide assistive spatialization, finding that orientation cues prior to unlocking do not help. Measurements of edit distance and dynamic time warping accuracy were collected, using a within-group, randomized study of 26 participants. 1021 passcode entry gestures were collected and classified, identifying six user strategies for using the pre-entry tactile feedback, and ten codes for types of events and errors that occurred during entry. We found that users who focused on orienting themselves to position the first digit of the passcode using the tactile feedback performed better in the task. These results could be applied to better define eyes-free behavior in further research, and to design better and more secure methods for eyes-free authentication.

Published by Elsevier Ltd.

1. Introduction

The threat of observational attacks in shared or public spaces may influence or modify the way smartphone users interact with their devices. In particular, users may favor unlocking their mobile devices out-of-view, without looking at the screen to avoid others from *surfing* the authenticator. Purposeful user obfuscation (e.g. keeping the screen out of sight from third parties or hidden cameras by hiding the device in the pocket or bag [1]) for purposes of the initial stages of the interaction, limits the likelihood of the authentication sequence being viewed. This can put users at some level of ease, even if the remainder of the interaction is performed in-view of third parties.

Eyes-free authentication behaviors may also be performed when the situation, context or environment demands it. For example, in situations where glare may be factor, or the environment is inappropriate for mobile device usage and discretion is needed (e.g. [2], the interaction may be performed away from view). While eyes-free interactions for different types of mobile device have been studied by researchers in the past [3–13], studies have yet to examine real world eyes-free authentication behaviors; investigating

the performance with common authentication mechanisms when the phone is out-of-view, and user coping strategies to enter passcodes in an eyes-free manner.

To address this knowledge gap, we conducted a randomized, multi-factor study with 26 participants entering PINs and gesture-based patterns (termed: "patterns" in this paper). Participants entered passcodes under both in-view and eyes-free conditions, as well as eyes-free using an additional training module for spatialization based on tactile feedback.

The tactile channel was chosen to discreetly offer cues directly to the user's hand, without drawing attention during interaction, as would likely occur with auditory or visual cues. Existing assistive aids aid to eyes-free PIN authentication, such as iOS VoiceOver, rely on audio feedback (audio readout of PIN number buttons when touched, allowing selection). However, audio cues impose usability and security penalties in shared and public spaces.

Biometric authentication such as fingerprint identification can greatly expedite this task for many users. However, fingerprint identification remains only a secondary means of authentication, which is generally tied to a PIN or patterns for screen unlocking. Essentially, even biometric authentication users must necessarily enter conventional passcodes on a semi-regular basis, and eyes-free conditions may apply in some instances.

In light of this, tactile-only feedback was designed for this study as a research device for understanding authentication performance

* Corresponding author.
  E-mail addresses: flynn.wolf@umbc.edu (F. Wolf), aviv@usna.edu (A.J. Aviv), rkuber@umbc.edu (R. Kuber).

with strictly eyes-free interaction. Its functionality, and our evaluation of its performance, is not intended to propose a workable real-world tool in the present form. Instead, we tried to capture how users develop techniques that use additional spatial cues to locate key screen features. This spatialization might then assist the accuracy and precision of eyes-free authentication gestures, especially for situations where the user may feel at risk of being a victim of an observer attack or be at risk of a situational impairment.

Given these assumptions, we have undertaken these research questions:

- *RQ1*: How well are users able to perform eyes-free authentication (without tactile feedback) with common methods, such as PIN and pattern entry, and how is this affected by the length and visual features of passcodes?
- *RQ2*: Will the relationship between spatial cues to screen layout features (e.g. position of buttons), presented by tactile interaction, enhance the user's performance when authenticating eyes-free?
- *RQ3*: When tactile feedback is presented, what approaches will users develop for using it?

With these considerations, during the experiment we collected complete movement traces, recording all participants' touch-based gestures during each authentication attempt, totaling 1021 eyes-free traces. To extend the work described in [14], we aimed to understand the input techniques and strategies the participants developed when completing the tasks. To do this, we classified all the traces, and developed a set of verified and grounded labels to describe the actions of the participants.

We further evaluated participants' performance in the eyes-free setting in two dimensions, accuracy and precision. For accuracy, we considered the edit-distance (or *Levenshtein distance*) between the input passcode and the true passcode. The edit-distance considers the number of additions or removals to transform one string sequence into another. For precision, we developed a geometric distance measure between in-view and eyes-free traces using *Dynamic Time Warping* (DTW), computing the average distance between temporally-associated points in the trace.

Based on this analysis, we found that participants using patterns were more accurate and precise in eyes-free settings, as compared to PINs. Additional tactile training was found to not improve the accuracy or precision of the participants' entries. We discuss users' observations regarding this distinction between task performances. When applying the classification results, we found that specific techniques in both training stages impacted performance. In particular, traces where participants used the additional tactile training aid to understand specifically the location of the starting digit of their passcode showed the most significant increase in performance, for both PINs and patterns. In addition to identifying techniques that improved performance, we also developed a set of classifications for eyes-free entry and training.

The results firstly contribute to an initial baseline of performance results and classifications of types for eyes-free interaction behaviors, events, and error types. We also show that the describe strategies for locating the starting location of authentication gestures (i.e. the screen position of the button for a passcode's first digit) that correspond with a number of significant effects on user performance. These results will help further research on eyes-free interaction make accurate comparisons and descriptions regarding this condition. Additionally, these insights will help iterate the design of targeted training aids for users, such as blind mobile technology users who rely on secure ubiquitous computing for privacy-sensitive tasks in shared spaces, who need to authenticate frequently in eyes-free settings (i.e. when at perceived risk of an observer attack described in [1]). Informing users of effective techniques will enable users to enter unlock authentication more confidently, securely, and accurately, away from adversarial observation.

While the tactile aid adopted for this study produced a mostly negative result from accuracy and edit distance measures, we assert several important contributions from this investigation:

1. A novel characterization of HCI and security performance conditions for eyes-free authentication tasks.
2. A systematic inquiry of accuracy, precision, and timing effects of input in eyes-free settings.
3. Establishing the unequivocal performance gap between eyes-free PIN and pattern entry (although unsurprising, this is the first time this has been shown empirically).
4. The extension of existing classification methodology for coding eyes-free unlock entry methods and events, similar to error codes established in von Zezschwitz et al. [15].
5. Identifying significant relationships between classification codes and authentication conditions (e.g. a decrease in Start-Hunt behavior for pattern passcodes ($\chi^2 = 8.17$, $p < 0.005$)).
6. Identifying passcodes features for which accuracy and/or precision significantly deviated from average (e.g. self-crossing pattern 743521).

We feel the relationship between the initial training methods that users develop using the tactile aid, such as those that help locate the starting point of the authentication gesture, are particularly illustrative. Strategies, such as the Start-Hunt trial code and Return to Start training code, offer an insight into the ways that users cope with the challenges of entering gestures under eyes-free conditions. By being able to better understand user strategies taken, along with events and error types made, this work could lead to the improved support of targeted training aids for users who interact with mobile authentication solutions under eyes-free conditions.

2. Related work

2.1. Eyes-free interaction techniques

As mobile technologies reduce in size and provide increasing amounts of PC-like functionality, these technologies become an attractive option for performing tasks while on-the-go. As information is predominantly presented via the graphical user interface, the user is heavily reliant on visual feedback to perform mobile tasks.

However, there are scenarios when difficulties are faced viewing the interface. One of the predominant issues relates to worries about third parties viewing content, and using this information without permission. Examples described by Yi et al. [16] include (1) environmental factors (e.g. excessive brightness impacting the user's ability to perceive screen content, and in scenarios where switching visual attention between the device and the physical environment poses safety concerns), (2) social factors (e.g. instances where it may be socially-inappropriate to view the screen, or multi-task in front of others), (3) constraints imposed by the mobile devices themselves (e.g. difficulties seeing content due to the crowded nature of content on mobile GUIs), and (4) personal factors (e.g. no perceived benefit to using vision to performing the task).

Additionally, if the user feels under threat of observer attacks, the screen may be hidden from view, either shielded by the hand [17], or placed within a garment or accessory [1]. The user can then attempt to use a combination of a mental image of the interface and muscle memory to attempt to interact with the device.

One of the fundamental motivations for eyes-free interaction is that as it leaves visual attention unoccupied, users are free to perform additional tasks [18]. However, performing mobile tasks

when visual and other forms of feedback are not available, can lead to errors during the input process, and contribute to levels of frustration among users. To better support users, a range of techniques have been developed involving gestural input (e.g. [3,5–7,11,13,19]), or voice input (e.g. [4,9]), along with accessible forms of output to provide feedback to the user (e.g. audio [5,6,10,13,19], and/or tactile output [8,19] either to the user's hand via the mobile device or via a separate wearable. Similar technologies have also been designed to support users for whom the visual channel is restricted or blocked (i.e., individuals who are blind and visually-impaired [20–24], with the aim of substituting or complementing visual feedback with other forms of accessible information). However, any solutions developed would ideally need to work in conjunction with existing assistive technologies (e.g., screen readers).

2.2. Augmenting interfaces to support eyes-free interactions

To help users orient position and better understand the layout of content on a mobile interface during eyes-free interactions, non-visual cues have been added to the interface. The PocketMenu solution [12] for pocket-based mobile device interactions, presents tactile information to convey the position and state of buttons, arrayed along the left edge of the mobile phone touchscreen, to exploit the natural tactile localization that the outer bezel of the phone affords. Findings from the researchers' study showed that PocketMenu outperformed auditory output (VoiceOver) in terms of completion time, selection errors, and subjective usability, making it ideal for interactions where the user is on-the-go [12]. Mc-Gookin et al. [25] used tactile overlays to support interaction with a visual touchscreen. Recommendations proposed to support eyes-free interactions include presenting a discernible button for orientation (i.e. home button), through the use of a tactile aid such as an adhesive bump-on, and providing feedback for all interactions. However, care should be taken with the latter, to ensure that the user is not overloaded with feedback. Oakley and Park [18] note the trade-off which designers should be aware of - between the amount of information contained within user interface feedback, the speed with which this can be achieved and the amount of effort and attention required to interpret it. This is thought to be especially important in the eyes-free domain. Existing research has not directly addressed tactile feedback using built-in actuators for eyes-free authentication, that would orient users to any passcode starting position.

2.3. Device and interaction modification for authentication

Although not designed specifically to improve accessibility of entry under eyes-free conditions, researchers have adopted a series of methods to reduce the likelihood of adversarial observations. Examples include the Back of Device scheme, proposed by De Luca et al. [26], who used a simulated double-sided touchscreen to allow users to grip the device in two hands, and conceal a secondary authentication action. This was found to be more resistant to simulated observation attacks than other common login methods. In terms of presenting information through alternative senses, the tactile channel has been a popular method of communicating information relating to the authentication discreetly to the user. Examples include the Haptic Wheel [27], where the user positions their hand around a rotary dial. After each input, the system randomizes the vibration it emits to protect the user from observer attack, and the VibraPass system [28], where the system presents tactile cues to the user's hands to indicate when to enter false PIN numbers to randomize each entry. While this modification added little time to a typical interaction (averaging 1.68 additional seconds), randomized PINs were intercepted by the simulated shoulder surfers 32.5% of the time (versus 100% for non-VibraPass pat-

Table 1
Conditions for study.

Eyes-free		In-view	
A) w/o tactile aid		B) w/o tactile aid	
A1) 4-digit PIN	A3) 4-len Pattern	B1) 4-digit PIN	B3) 4-len Pattern
A2) 6-digit PIN	A4) 6-len Pattern	B2) 6-digit PIN	B4) 6-len Pattern
C) w/ tactile aid		D) w/ tactile aid	
C1) 4-digit PIN	C3) 4-len Pattern	D1) 4-digit PIN	D3) 4-len Pattern
C2) 6-digit PIN	C4) 6-len Pattern	D2) 6-digit PIN	D4) 6-len Pattern

terns). Other tactile solutions include H4Plock, proposed by Ali et al. [29], where the user is required to enter a sequence of up to four pre-selected on-screen gestures while responding to tactile prompts signaling whether stimuli from a primary or secondary passcode should be entered. The solution proved to be secure against 76.5% of participants, who carried out attacks immediately after watching a set of videos. Participants were able to express strong levels of confidence in using the system.

Limited work has been conducted exploring existing common mobile authentication mechanisms and their use when the user needs to purposely obfuscate the screen to perform tasks in an eyes-free manner. In this paper, we describe a study investigating PIN and graphical pattern entry using tactile feedback, when the device is out of view, with the aim to unlocking entry. While auditory feedback appears to be an appealing solution to this scenario, it may be impractical. Providing an auditory representation of screen content may be insecure or an unacceptable distraction. Furthermore, auditory cues from a mobile device may be masked by ambient noise, which may pose challenges when attempting to authenticate entry. Tactile feedback may offer a solution to directing users to make accurate entry gestures (e.g. finding the start position of their passcode). A tactile aid to support orientation has also been evaluated, as part of this research. To better understand user behaviors, we aim to focus on input techniques, and methods of classifying these.

3. Study design and procedure

We designed a within-subjects, multi-factor study where participants were asked to enter authentication sequences under four primary conditions. To address RQ1, out-of-view performance was measured and compared to that of in-view. RQ2 is addressed by the inclusion of with- and without-tactile feedback conditions. The conditions are described in more detail below:

- *in-view w/o training aid*: participants performed the tasks completely in-view where they could view the device and the entry thereon, without a training aid.
- *in-view w/ training aid*: participants performed the tasks completely in-view with a tactile/vibration-based training aid.
- *eyes-free w/o training aid*: participants performed the tasks out-of-view (eyes-free) without a training aid.
- *eyes-free w/ training aid*: participants performed the task with a tactile/vibration-based training aid.

Within each condition, the participants were assigned a sequence of 10 PINs or patterns to enter, of which 5 were length 4 PINs/patterns and 5 were length 6 PINs/patterns. We summarize these conditions and the set of authentication passcodes in Tables 1 and 2, respectively. Passcodes were sourced from real world data to establish validity, and selected individually to include important visual characteristics for analysis, for example left vs. right side shift, and self crossing patterns. For all experiments, we used a Nexus 5x, which has a common 5.79″ x 2.86″ form factor,5.2″ display, 1080 x 1920 resolution. The steps of the study are illustrated in Fig. 1. The procedure and study design are similar to that presented in related work [30,31].
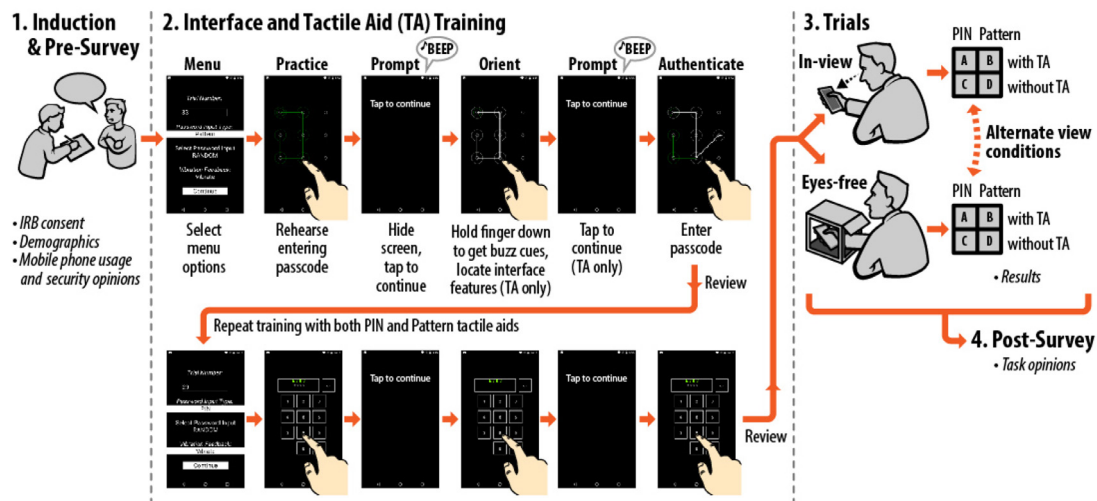
Fig. 1. Study procedure: Participants are informed of the make-up of the experiment, and basic demographic and usage data is collected in step (1). In step (2), participants are trained on the application interface for data collection for both PINs and patterns, and spend time interacting with the training aid (the first prompt/orient phase). Without the training aid, participants in trial A and B proceed directly to the next phase. After completing the training stage, step (3) the trials begin alternating between the conditions as presented in Table 1.

Table 2
PINs and patterns used in experiments, and reference for labeling pattern contact points. See the Appendix and Supplementary Materials for graphical depictions of the authentication.

	PINs	Patterns	
4-length	1328	0145	
	1955	1346	
	5962	3157	
	6702	4572	
	7272	6745	
6-length	153525	014763	
	159428	136785	
	366792	642580	
	441791	743521	
	458090	841257	

Pattern grid reference:
```
 0   1   2
 3   4   5
 6   7   8
```
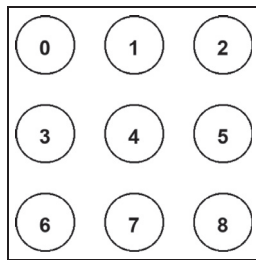
3.1. Study procedure

As presented in Fig. 1, the study consists of three steps: (1) introduction/pre-study, (2) training stage, and (3) experimental trials. In total, the procedure took about 40 min per participant, as training time was allocated to familiarize participants with relevant authentication issues, the two training aids, and the methods of the study. This was undertaken to prepare participants at least up to a low level of performance that could simulate everyday usage. To begin, in step (1), participants provided informed consent of participation (as required by the IRB), and were asked a series of pre-survey questions. This included demographic questions as well as questions regarding their mobile phone usage and locking habits.

The training phase followed. As we developed a specialized data collection application that directed and stepped users through the procedure (see Fig. 1 for more details), it was necessary to provide training and familiarization with the interface to allow participants to focus on accuracy and precision using the tactile feedback rather than any unnecessary memorization tasks. In particular, we wished for participants to be fully aware of each of the authentication codes (i.e., all the PIN and pattern examples that they would encounter so they would be familiar with them when entering them in an eyes-free manner). A visual reference

of the passcode to be entered was always available throughout all conditions, via a laptop showing the passcode. We also wished to train participants on the tactile aid that some of the trials would employ (condition C and D in Table 1), so that they could maximize the use of the aid. We also wanted to give participants some simulated trial runs under eyes-free conditions, where they could subsequently review their performance by viewing replays of their entries, provided within the data collection application. Once participants were comfortable with the interface, the procedure, and the tactile feedback aid, the trials for each condition began. The conditions were randomized using a Latin Square.

3.2. Interface and tactile feedback design

A web-based application was developed using HTML5 to collect data for the study. When opened in a smartphone's web browser, the application interfaces closely simulated the layout, font, animation, luminance, color, and interaction of typical mobile authentication screens. The pattern interface displayed the standard Android OS 3x3 grid, and the PIN interface displayed the standard 0–9 digit layout (telephone dial-pad style), including a text display that showed entered digits (which are obfuscated to an "*" after 1 s), a back/delete button next to the display, and an "OK" button to the right of the "0" to submit the PIN. Graphical depictions of the application are presented in Fig. 1 with the study's procedures.

The data collection applications interface is designed to assist the study procedure by presenting the participant with an authentication code and then directing users to place the device in the eyes-free setting to complete the experiment. Additionally, to support RQ3, the interface implemented the tactile feedback aid if the condition called for that. The interface application also collected and stored traces of all touch interactions. A trace consisted of the x- and y-coordinate of a touch event and the time in which that event occurred. To direct participants in the eyes-free setup, the interface played audible "beeps" to indicate to the participants transition between the steps of the experiment.

To facilitate the eyes-free setting, we used a shielding box in which the participant can interact with and hold the device in a natural posture, but cannot see the device itself (see Fig. 2). Cutouts in the sides of the shielding box allowed the researchers to observe the interaction, which the participant is unable to view. We placed no restrictions on how the user chose to interact with

Fig. 2. Participant holding the LG Nexus 5x test phone inside the eyes-free observation box.

Table 3
Demographics of participants.

		Male	Female	Total
	total	12	14	26
age	18–24	7	7	14
	24–34	5	7	12
OS	Android	8	5	13
	iOs	4	9	13
Unlock choice	Fingerprint	5	10	15
	PIN-6	2	7	9
	PIN-4	5	4	9
	Pattern	4	2	9
	No-lock	1	1	2
Level of confidence (1–5 Likert)	Phone security	3.75	3.79	3.77
		(STD: 1.42)	(STD: 0.7)	(STD: 1.07)
	Shoulder surfing	3.08	3.79	3.46
		(STD: 1.24)	(STD: 0.97)	(STD: 1.140

the device, i.e, using one or two hands, as long as they were not able to view the device directly.

The interface also incorporated tactile feedback in the form of vibrations for touches that occurred within a digit/point. This feature was not originally included in our experiment, and was added after prototyping the procedures based on participant feedback. We believe this is a reasonable choice given that most unlock authentication systems already incorporate this style of feedback.

In the conditions with the tactile training aid, participants were able to use the aid to orient with the interface prior to beginning the eyes-free authentication stage. The orientation occurred during continual touching, where a participant can swipe around the screen, receiving tactile/vibrational feedback when swiping over a digit/point. If the digit/point is the first digit/point in the PIN/pattern, a vibration cue encoded with a faster rhythm was used to differentiate this point from the others; however, no other passcode specific feedback was provided. Once the participant lifts their finger, ending the continuous swiping, the interface "beeps" indicating they should attempt to enter the authentication. Trace information was collected during tactile training phase as well in the same format as in the authentication phase.

3.3. Pre- and post-Survey procedures

In steps (1) and (3), we asked participants a set of pre- and post-survey questions. Copies of the questions are presented in the Appendix. The questions included basic demographic information, as well as inquiries regarding mobile device usage, such as OS, time, and locking behavior. Reported statistics are available in Table 3.

In the pre-survey portion, participants were also asked a series of questions about how they chose passcodes, their concern for phone security, experiences with device theft or unauthorized access, and how those factors might have affected their behavior. Two questions were Likert responses, on a scale from 1 to 5:

- *How concerned are you with keeping your phone secure (1, not at all concerned, to 5, highly concerned)?*
- *How concerned are you typically, in public spaces, with the threat of someone watching you authenticate and collecting your passcodes (1, not at all concerned, to 5, highly concerned).*

Responses to these questions did not significantly correlate with performance on the task, but did provide some information on the

participants usage and locking behavior and mindset that is relevant to the task. The responses are also presented in Table 3.

Following the tasks, during the post-survey, we asked questions regarding the ease or difficulty of the task itself, with or without the tactile feedback aid. These results are presented in the results section of the paper.

3.4. Passcode selection

To select the passcodes, both PINs and patterns, we sought to use a representative sample from real world data for validity, that also had good spatial features in terms of the regions of the screen that must be touched. As such, we used the set of PINs and patterns published in related work [32,33]. The set of passcodes are presented in Table 2 and visuals are provided in the Appendix and the Supplementary Materials [1].

Following [32], the PINs from the study were culled from publicly leaked password sets, using sequences of 4 or 6 digits. This is an acceptable practice in simulating PINs from related work [34], as good realistic data is not available for purposes of research. Patterns from [32] were selected from [35] using spatial criteria, where the authors ensured good spatial criteria: up, down, left, right, and neutral shifts.

3.5. Participants and recruitment

Participants were recruited via a university mailing list, alongside the use of the snowball sampling technique. The latter was used with the aim of recruiting a wider, more representative group of mobile device users. There were 26 individuals recruited, comprising of 12 male and 14 female younger adults. Participants were fully sighted mobile users, evenly divided between using iOS and Android. For their current unlock conditions, many used their fingerprint in addition to another lock mechanism, such as a PIN or pattern. Only 2 participants had no lock authentication on their smartphone. Details are provided in Table 3.

3.6. Realism and limitations

As the study was performed under controlled in-person conditions, the set-up introduced some constraints which reduced the level of realism. Examples include the slightly extended posture of the users arm, holding and pointing with the mobile device when interacting within the shielding box. Further work will examine more natural methods of interacting with the device under more

---

realistic constraints (e.g. the impact of the user entering passcodes while the device is located in his/her pocket or bag).

In order to provide a baseline control, participants were asked to perform in-view PIN and pattern conditions, with/without the presence of a tactile aid. Due to the randomization of conditions, half of the group of participants performed this condition after the eyes-free condition. This was conducted to minimize the likelihood of an order effect. It was acknowledged that this may have led to a slight performance disadvantage for users performing eyes-free first, as the in-view conditions may be considered equivalent to a small amount of extra training.

In terms of limitations, passcodes were presented in the same order to all participants under each condition. Although passcodes did not necessarily increase in complexity during each condition, it is acknowledged that this may have contributed to an effect.

While the number of participants recruited for our study either exceeded or remained in line with those selected for other studies relating to eyes-free interaction (e.g. [12,25]), the sample selected for the study described in this paper, can be considered small in size (n=26), and limited in range of demographic features. While significance was found within our results, challenges can be faced when samples are very small, which can impact the power of the study and increase the margin of error. Similarly, the representativeness of findings may be impacted if the samples are too limited in terms of age. The sample in our study reflected the university environment (i.e. younger adults). It is also acknowledged that participants who were primarily iOS users may have been at a slight disadvantage in levels of experience interacting with stroke-based patterns, as these types of mechanism are found on devices running the Android OS. However, training was provided to familiarize users with the process, which helped to offset any potential disparity. Finally, as participants came from a university, issues of bias may have also been introduced from the working environment. While it is not uncommon in authentication studies to recruit participants comprising mainly of students, future studies would aim to widen the pool, with the aim of recruiting a more diverse sample, who are more representative of the array of individuals who use mobile devices (e.g., participants with varying levels of concern about security).

4. Results

In this section, we present the results of both the performance of the participants, as well as the classification of the traces and training techniques. We begin with a description of the metrics applied, followed by the performance results, and finish with the classification results.

4.1. Performance metrics

Accuracy. A crucial and informative metric to determine effectiveness is simply how accurately participants performed the tasks in eyes-free settings. Of course, authentication is a binary response: either participants entered the passcodes correctly or not. As we are also interested in granularity of performance, we also considered the *edit distance*, normalized to the distance of each passcode. More precisely, we considered accuracy as a fraction calculated for a passcode p and entered code p'

$$acc = \frac{len(p) - d(p, p')}{len(p)}$$

where $len(p)$ is the length of the passcode and $d(p, p')$ is the edit distance between the entered and expected passcode. The edit distance (or Levenshtein distance) computes the number of additions, subtractions, or replacements needed to transform one sequence into another. For example, if the task requires entering the passcode 123456, and the participant entered 12356 (or any passcode off by one in some dimension), then the accuracy would be $(6 - 1)/6$ or 0.83 as the edit distance is 1. This is a generous accuracy measure in the sense that the edit distance is a greedy algorithm and tries to aggressively match strings. However, given the nature of the task, eyes-free entry, we feel that this provides a better reflection on participant effort and performance than a binary yes/no.

Precision. As a second measure of performance, we wish to compare the traces of the eyes-free entry to that of the in-view entry. Recall that interface application collects detailed trace information for each touch event in the form of $d = (x, y, t)$ where x is a width coordinate, y is a height coordinate, and t is a time indication from the start of the interaction. The goal is to develop a method that allows two sequences, which may be time dilated with different number of points, to be compared in the 2-D space.

The solution to this was the use of Dynamic Time Warping (DTW), which has been used extensively in the space of free-form gesture authentication [26,36–38]. DTW takes two time series sequences and aligns one to the other to best match the time dilation present. For example, in Fig. 3(a), the blue trace is the eyes-free trace and the red trace is the in-view trace. Notice that the blue trace (eyes-free) is longer in time (will have more points in the trace) and also have more loops due to mistakes during entry. Comparing these two traces directly, such as point by point, would indicate that the participant performed the authentication task quite poorly, but in fact, here, this is an 100% accurate entry of pattern 136785.

With DTW, we can better show the precision of the user. First, DTW will associate points in one time series with the other based on a distance measure, which is Euclidean in this case. In Fig. 3(a), this matching is represented by the green lines. Based on the matching, point reduction is performed by averaging the points in the blue trace that match a single point in the red trace, giving us Fig. 3(b). The last step in the metric is to take the average distance between two matched points, again using Euclidian, which represents the precision of the trace in the eyes-free setting to the in-view one.

We treated PIN traces as a series of distinct touch event traces since each touch could really be a drag/swipe and consist of multiple points in the trace. To apply the DTW distance metric for precision, we merged the touch event traces into a single trace, via concatenation, and then applied the same routine as described previously. This resulted in extra connections when the participant transitioned between touch events, but this information is useful when considering the precision as the direction and regularity of those transitions are measurable and meaningful. We excluded the last touch event from the analysis, pressing "OK" as this was not always successful in the eyes-free conditions, and the researcher monitoring the experiment would need to assist in this process.

4.2. Performance results

The primary performance results are presented in Table 4 with per passcode breakdowns in Tables 5 and 6. We applied the two metrics (accuracy and precision) described previously to conditions that considered the use or non-use of the tactile aid with the two unlock authentications, PINs and patterns. Additionally, we considered two further measurements, the precision of just the start point, calculated as the Euclidean distance between the first touch event in the in-view to the eyes-free setting, and the time (in milliseconds) of entering the authentication in the eyes-free setting.

As observed in Table 4, the impact of the tactile aid is rather limited. There were small effects for patterns, as compared to no effect in PINs, and the effect was most notable when it comes
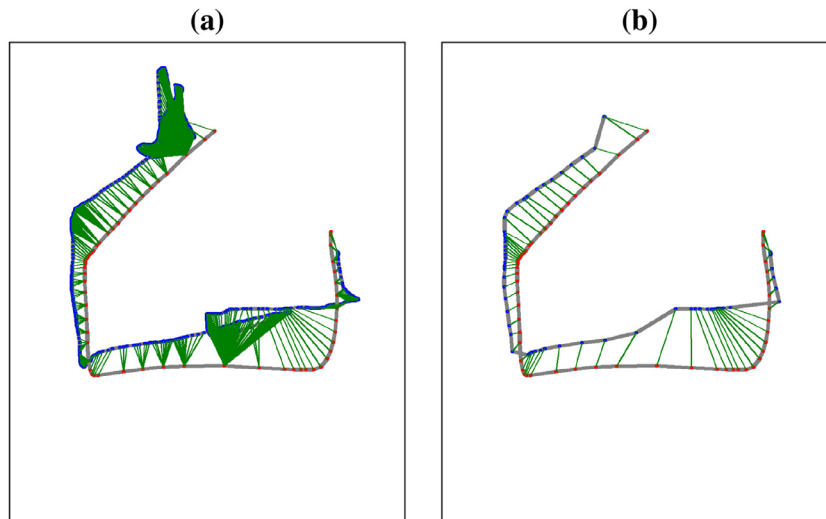
(a) **(b)**

Fig. 3. Example of using DTW to do point reduction: In (a) The blue trace is eyes-free and the red trace is in-view, and the green lines show the points in the blue trace time sequence matched to points in the red trace; and in (b), associated points in the blue trace are merged via Euclidean average to form a set of singly matched points to the red trace. The precision metric is then the average Euclidean distance between the matched point of the blue trace with the red one. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Performance results: _acc._ is the accuracy using the edit-distance measure, _prec._ is the precisions using the DTW method, `start` is the Euclidian distance of the start point, and `time` refers to the number of milliseconds. As the data followed a normal distribution, we used a two-tailed _t_-test. Horizontally, the _t_-test compared Pattern vs. PIN results, and vertically, the _t_-test compared w/ and w/o the tactile aid. Effect size of $\alpha = 0.05$ was considered significant. Only traces that were complete and collected without errors were considered.

	Tactile Aid	Pattern	PIN	_t_-test
acc.	w/o	($\mu = 0.80$, $\sigma = 0.28$, $n = 250$)	($\mu = 0.72$, $\sigma = 0.30$, $n = 260$)	$t = 3.08$, $p < 0.05^*$
	w/	($\mu = 0.83$, $\sigma = 0.27$, $n = 253$)	($\mu = 0.71$, $\sigma = 0.30$, $n = 258$)	$t = 4.78$, $p < 0.001^{**}$
	_t_-test	$t = 1.17$, $p = 0.243$	$t = -0.49$, $p = 0.628$	
prec.	w/o	($\mu = 80.93$, $\sigma = 60.40$, $n = 250$)	($\mu = 145.18$, $\sigma = 90.29$, $n = 260$)	$t = -9.41$, $p < 0.001^{**}$
	w/	($\mu = 83.21$, $\sigma = 69.74$, $n = 253$)	($\mu = 154.49$, $\sigma = 94.11$, $n = 258$)	$t = -9.71$, $p < 0.001^{**}$
	_t_-test	$t = 0.39$, $p = 0.695$	$t = 1.15$, $p = 0.251$	
start	w/o	($\mu = 80.51$, $\sigma = 67.62$, $n = 250$)	($\mu = 96.56$, $\sigma = 109.79$, $n = 260$)	$t = -1.98$, $p < 0.05^*$
	w/	($\mu = 68.35$, $\sigma = 61.96$, $n = 253$)	($\mu = 91.77$, $\sigma = 86.46$, $n = 258$)	$t = -3.51$, $p < 0.001^{**}$
	_t_-test	$t = -2.10$, $p < 0.05^*$	$t = -0.55$, $p = 0.582$	
time (ms)	w/o	($\mu = 7134.86$, $\sigma = 6672.95$, $n = 250$)	($\mu = 8597.35$, $\sigma = 6314.62$, $n = 260$)	$t = -2.54$, $p < 0.05^*$
	w/	($\mu = 7500.12$, $\sigma = 5359.63$, $n = 253$)	($\mu = 10227.07$, $\sigma = 8236.91$, $n = 258$)	$t = -4.43$, $p < 0.001^{**}$
	_t_-test	$t = 0.68$, $p = 0.499$	$t = 2.53$, $p < 0.05^*$	

Table 5
Performance Metrics per-Pattern: Comparisons were made with-out (w/o) and with (w/) the tactile aid, and a _t_-test is used as the data is normal. Note that not all participants provided valid traces, and invalid traces were excluded.

Pattern	Metric	w/o Tactile aid	w/ Tactile aid	_t_-test
0145	_acc_	($\mu = 0.81$, $\sigma = 0.27$, $n = 25$)	($\mu = 0.86$, $\sigma = 0.27$, $n = 25$)	$t = -0.66$, $p = 0.510$
	prec	($\mu = 82.66$, $\sigma = 50.16$, $n = 25$)	($\mu = 71.77$, $\sigma = 62.11$, $n = 25$)	$t = 0.68$, $p = 0.499$
1346	_acc_	($\mu = 0.88$, $\sigma = 0.19$, $n = 25$)	($\mu = 0.87$, $\sigma = 0.22$, $n = 26$)	$t = 0.25$, $p = 0.801$
	prec	($\mu = 57.19$, $\sigma = 28.84$, $n = 25$)	($\mu = 72.96$, $\sigma = 37.22$, $n = 26$)	$t = -1.69$, $p = 0.098$
3157	_acc_	($\mu = 0.78$, $\sigma = 0.31$, $n = 25$)	($\mu = 0.75$, $\sigma = 0.28$, $n = 26$)	$t = 0.36$, $p = 0.717$
	prec	($\mu = 89.60$, $\sigma = 77.38$, $n = 25$)	($\mu = 87.02$, $\sigma = 45.19$, $n = 26$)	$t = 0.15$, $p = 0.885$
4572	_acc_	($\mu = 0.92$, $\sigma = 0.14$, $n = 25$)	($\mu = 0.80$, $\sigma = 0.29$, $n = 25$)	$t = 1.86$, $p = 0.068$
	prec	($\mu = 67.62$, $\sigma = 31.71$, $n = 25$)	($\mu = 130.00$, $\sigma = 125.20$, $n = 25$)	$t = -2.41$, $p = 0.020^*$
6745	_acc_	($\mu = 0.68$, $\sigma = 0.38$, $n = 25$)	($\mu = 0.74$, $\sigma = 0.42$, $n = 25$)	$t = -0.53$, $p = 0.595$
	prec	($\mu = 111.61$, $\sigma = 96.17$, $n = 25$)	($\mu = 90.51$, $\sigma = 80.56$, $n = 25$)	$t = 0.84$, $p = 0.405$
014763	_acc_	($\mu = 0.81$, $\sigma = 0.26$, $n = 25$)	($\mu = 0.87$, $\sigma = 0.24$, $n = 26$)	$t = -0.74$, $p = 0.461$
	prec	($\mu = 82.28$, $\sigma = 68.34$, $n = 25$)	($\mu = 73.37$, $\sigma = 66.77$, $n = 26$)	$t = 0.47$, $p = 0.640$
136785	_acc_	($\mu = 0.80$, $\sigma = 0.23$, $n = 25$)	($\mu = 0.86$, $\sigma = 0.22$, $n = 25$)	$t = -0.93$, $p = 0.356$
	prec	($\mu = 70.12$, $\sigma = 37.28$, $n = 25$)	($\mu = 80.78$, $\sigma = 53.51$, $n = 25$)	$t = -0.82$, $p = 0.418$
642580	_acc_	($\mu = 0.78$, $\sigma = 0.25$, $n = 25$)	($\mu = 0.79$, $\sigma = 0.23$, $n = 25$)	$t = -0.20$, $p = 0.844$
	prec	($\mu = 78.69$, $\sigma = 40.08$, $n = 25$)	($\mu = 109.83$, $\sigma = 76.32$, $n = 25$)	$t = -1.81$, $p = 0.077$
743521	_acc_	($\mu = 0.77$, $\sigma = 0.27$, $n = 25$)	($\mu = 0.92$, $\sigma = 0.17$, $n = 26$)	$t = -2.34$, $p = 0.024^*$
	prec	($\mu = 89.66$, $\sigma = 69.72$, $n = 25$)	($\mu = 54.70$, $\sigma = 26.51$, $n = 26$)	$t = 2.38$, $p = 0.021^*$
841257	_acc_	($\mu = 0.80$, $\sigma = 0.31$, $n = 25$)	($\mu = 0.86$, $\sigma = 0.22$, $n = 24$)	$t = -0.79$, $p = 0.434$
	prec	($\mu = 79.91$, $\sigma = 49.08$, $n = 25$)	($\mu = 62.14$, $\sigma = 32.16$, $n = 24$)	$t = 1.49$, $p = 0.142$

Table 6

Performance Metrics per-PIN: Comparisons were made with-out (w/o) and with (w/) the tactile aid, and a *t*-test is used as the data is normal. Note that not all participants provided valid traces, and invalid traces were excluded.

PIN	Metric	w/o Tactile aid	w/ Tactile aid	*t*-test
1328	*acc*	($\mu = 0.73$, $\sigma = 0.32$, $n = 26$)	($\mu = 0.69$, $\sigma = 0.33$, $n = 26$)	$t = 0.43$, $p = 0.672$
	prec	($\mu = 110.42$, $\sigma = 54.18$, $n = 26$)	($\mu = 108.75$, $\sigma = 41.15$, $n = 26$)	$t = 0.12$, $p = 0.901$
1935	*acc*	($\mu = 0.73$, $\sigma = 0.36$, $n = 26$)	($\mu = 0.62$, $\sigma = 0.36$, $n = 26$)	$t = 1.06$, $p = 0.296$
	prec	($\mu = 149.54$, $\sigma = 125.19$, $n = 26$)	($\mu = 169.98$, $\sigma = 104.29$, $n = 26$)	$t = -0.64$, $p = 0.525$
5962	*acc*	($\mu = 0.82$, $\sigma = 0.20$, $n = 26$)	($\mu = 0.72$, $\sigma = 0.40$, $n = 26$)	$t = 1.09$, $p = 0.280$
	prec	($\mu = 136.19$, $\sigma = 81.36$, $n = 26$)	($\mu = 166.62$, $\sigma = 119.08$, $n = 26$)	$t = -1.08$, $p = 0.287$
6702	*acc*	($\mu = 0.65$, $\sigma = 0.29$, $n = 26$)	($\mu = 0.63$, $\sigma = 0.28$, $n = 26$)	$t = 0.24$, $p = 0.810$
	prec	($\mu = 176.51$, $\sigma = 98.64$, $n = 26$)	($\mu = 172.79$, $\sigma = 97.10$, $n = 26$)	$t = 0.14$, $p = 0.891$
7272	*acc*	($\mu = 0.81$, $\sigma = 0.31$, $n = 26$)	($\mu = 0.67$, $\sigma = 0.31$, $n = 26$)	$t = 1.56$, $p = 0.124$
	prec	($\mu = 174.64$, $\sigma = 112.77$, $n = 26$)	($\mu = 176.16$, $\sigma = 100.48$, $n = 26$)	$t = -0.05$, $p = 0.959$
153525	*acc*	($\mu = 0.78$, $\sigma = 0.28$, $n = 26$)	($\mu = 0.82$, $\sigma = 0.21$, $n = 26$)	$t = -0.66$, $p = 0.515$
	prec	($\mu = 115.73$, $\sigma = 62.08$, $n = 26$)	($\mu = 136.91$, $\sigma = 78.55$, $n = 26$)	$t = -1.08$, $p = 0.286$
159428	*acc*	($\mu = 0.69$, $\sigma = 0.26$, $n = 26$)	($\mu = 0.72$, $\sigma = 0.25$, $n = 26$)	$t = -0.45$, $p = 0.654$
	prec	($\mu = 125.48$, $\sigma = 53.41$, $n = 26$)	($\mu = 144.00$, $\sigma = 92.45$, $n = 26$)	$t = -0.88$, $p = 0.380$
366792	*acc*	($\mu = 0.66$, $\sigma = 0.32$, $n = 26$)	($\mu = 0.71$, $\sigma = 0.24$, $n = 26$)	$t = -0.65$, $p = 0.517$
	prec	($\mu = 178.64$, $\sigma = 99.21$, $n = 26$)	($\mu = 196.46$, $\sigma = 107.88$, $n = 26$)	$t = -0.62$, $p = 0.538$
441791	*acc*	($\mu = 0.72$, $\sigma = 0.27$, $n = 26$)	($\mu = 0.81$, $\sigma = 0.17$, $n = 24$)	$t = -1.46$, $p = 0.151$
	prec	($\mu = 173.36$, $\sigma = 84.65$, $n = 26$)	($\mu = 163.50$, $\sigma = 73.30$, $n = 24$)	$t = 0.44$, $p = 0.663$
458090	*acc*	($\mu = 0.65$, $\sigma = 0.32$, $n = 26$)	($\mu = 0.71$, $\sigma = 0.28$, $n = 26$)	$t = -0.62$, $p = 0.540$
	prec	($\mu = 111.25$, $\sigma = 51.50$, $n = 26$)	($\mu = 110.44$, $\sigma = 53.00$, $n = 26$)	$t = 0.06$, $p = 0.956$

to the start point in the pattern. Using the aid showed a significant improvement in starting accuracy as compared to not using the tactile aid. This is reasonable given that the tactile aid performed different vibration feedback for the first point/digit in the passcode. As the *t*-test performed is two-tailed, one can consider the significance for the accuracy measure is somewhat intriguing, $p = 0.243/2 = 0.1215$. While this effect is not-significant as a one-tailed result, it is encouraging for tactile aids in the eyes-free setting, perhaps a better design based on feedback from this study, could improve the accuracy of entry.

The performance of the tactile aid for PINs can only be explained as detrimental. The eyes-free task for PINs is already significantly harder, in all conditions, but the addition of the aid showed worse performance for precision with no effect on accuracy (maybe even hurting accuracy). The time increase of using the aid is also striking and much larger than the time increase for patterns. We can speculate as to the reason for this disparity from participants post hoc responses regarding eyes-free PIN entry. It may be that the task is already challenging enough (concentrating on a series of discrete gestures hitting PIN digits) that the cognitive burden of integrating the aids tactile feedback only decreased the participants abilities. This is an area of future investigation, suggesting that different kinds of tactile aids may be needed for different authentication systems in the eyes-free setting.

When observing performance impacts of the tactile aid on a per-passcode basis, see Tables 5 and 6, there is no noticeable effect for either accuracy or precision. In only one case, for pattern 743521, was there a significant difference gained from using the tactile aid for both accuracy and precision. This pattern required doubling-back, (see the Supplemental Material for a graphical depiction) and the tactile aid may have provided some reference points for that process which improved performance. However, there were some negative results associated with using the tactile aid. For pattern 4572, there was an effect observed for precision, as the tactile aid *increased* the precision score, which degrades the precision of the pattern (smaller distances are better). This may have been the case that due to the compactness of this pattern, the tactile aid encouraged participants to take more elaborate/longer paths than they would have normally. In either case, this is further evidence that introducing aids for the eyes-free setting needs further investigation.

4.3. Classification metrics

As a further metric for understanding the techniques that participants applied to eyes-free authentication, we developed a set of classification labels over both the authentication task and the tactile training. These labels describe observed types of errors (e.g. Transposed or Missed OK) or actions (e.g. Went Well or Start Hunt) made during passcode entry by participants, or apparent strategies used to spatialize using the tactile aid (e.g. Wide Explorer or Start Point Only). For consistency in label development, we adopted a similar approach to that of Micinski et al. [39] and von Zezschwitz et al. [15]. Three researchers independently classified sub-sets, then met periodically to resolve differences between categorization labels. Although the definitions of some labels were mutually exclusive (e.g. a gesture could not be both Went Well and also include a Start Middle or End Hunt event), multiple compatible labels were applied to completely describe each gesture. Once agreement was reached on labels, one researcher labeled all the data while the other labeled a 15% random sample. Comparing the ratings labels, there was strong agreement between the two researchers using Cohens *kappa* ($\kappa = 0.900$, $p < 0.0005$). Additionally, during this procedure the researchers marked traces as invalid any observed miss-touches or other errors, which were then removed from the data sets.

The procedure produced six training labels and ten authentication labels. Visual examples of these labels are provided in Figs. 4 and 5 and descriptions of the labels, with counts, are presented in Tables 7 and 8. Animated examples of the classifications are provided on the supplementary web site [2].

4.4. Classification results

The frequency results for the different classification labels are presented in Tables 7 and 8. The frequencies are calculated by the total traces in the categories that were labeled as valid. Some traces were excluded due to either collection errors or misunderstandings on the part of the participant.

For the training classifications (see Table 7), the most common method for entering PINs was to perform a partial exploration of

---

[2] Supplementary web site - https://unlockclassification.wordpress.com

Table 7

Descriptions of the Training Classifications and Frequencies: Frequencies are calculated based on the total number of traces that received the label. Note that the trace can receive multiple labels, for example, it is possible to trace the path and return to the start. There was a total of 496 training traces considered, 248 pattern traces and 248 PIN traces.

Classification name	Description	PIN	Pattern
Random touch	Short gesture away from path	4 (1.61%)	0 (0.00%)
Partial explore	Touched part of space or path	100 (40.32%)	57 (22.98%)
Start point only	Only traced the area around the first digit	59 (23.79%)	47 (18.95%)
Return to start	Returns to starting point at end of gesture	37 (14.92%)	37 (14.92%)
Wide explore	Touched most of the space	32 (12.90%)	23 (9.27%)
Traced path	Traced the path of the entire passcode	54 (21.77%)	119 (47.98%)

Table 8

Descriptions of the Entry Classifications: Frequencies are calculated based on the total number of traces that received that label. Note that the trace can receive multiple labels, for example, it is possible to trace the path and return to the start. There was a total of 993 entry traces considered, 490 pattern traces and 503 PIN traces.

Classification name	Description	PIN		Pattern	
		w/o Tactile aid	w/ Tactile aid	w/o Tactile aid	w/ Tactile aid
Start-hunt	Missed or hunted for 1st digit/point	61 (12.13%)	49 (9.74%)	66 (13.47%)	37 (7.55%)
End-hunt	Missed or hunted for last digit/point	77 (15.31%)	69 (13.72%)	64 (13.06%)	66 (13.47%)
Middle-hunt	Missed or hunted for midrange digits	104 (20.68%)	116 (23.06%)	95 (19.39%)	103 (21.02%)
Went well	No major errors	111 (22.07%)	94 (18.69%)	105 (21.43%)	117 (23.88%)
Gave up	Just gave up and moved randomly	2 (0.40%)	1 (0.20%)	2 (0.41%)	6 (1.22%)
Transposed	Entered right shape in wrong place	11 (2.19%)	3 (0.60%)	12 (2.45%)	3 (0.61%)
Swift	Moved swiftly	1 (0.20%)	5 (0.99%)	1 (0.20%)	1 (0.20%)
Hesitant	Moved hesitantly	2 (0.40%)	1 (0.20%)	1 (0.20%)	0 (0.00%)
Used delete (PIN-only)	Used Delete button to back up	11 (2.19%)	8 (1.59%)	–	–
Missed ok (PIN-only)	Tried to tap OK but missed > 2–3 times	43 (8.55%)	43 (8.55%)	–	–

Table 9

Performance Metrics for Training Classifications: Use Kruskal's H-test to test for significance.

Classification	PIN		Pattern	
	acc.	prec	acc.	prec.
Random touch	($\mu = 0.69$, $\sigma = 0.32$)	($\mu = 142.75$, $\sigma = 140.53$)	–	–
Partial explore	($\mu = 0.73$, $\sigma = 0.29$)	($\mu = 139.14$, $\sigma = 78.03$)	($\mu = 0.78$, $\sigma = 0.27$)	($\mu = 84.08$, $\sigma = 47.74$)
Start point only	($\mu = 0.78$, $\sigma = 0.31$)	($\mu = 126.27$, $\sigma = 77.20$)	($\mu = 0.88$, $\sigma = 0.18$)	($\mu = 75.29$, $\sigma = 57.68$)
Return to start	($\mu = 0.76$, $\sigma = 0.24$)	($\mu = 124.19$, $\sigma = 59.80$)	($\mu = 0.91$, $\sigma = 0.22$)	($\mu = 66.41$, $\sigma = 53.93$)
Wide explore	($\mu = 0.69$, $\sigma = 0.31$)	($\mu = 194.93$, $\sigma = 114.63$)	($\mu = 0.80$, $\sigma = 0.38$)	($\mu = 100.48$, $\sigma = 94.65$)
Traced path	($\mu = 0.67$, $\sigma = 0.25$)	($\mu = 169.40$, $\sigma = 88.19$)	($\mu = 0.86$, $\sigma = 0.26$)	($\mu = 75.51$, $\sigma = 57.97$)
	$H = 8.268$, $p = 0.14$	$H = 17.74$, $p < 0.005^{**}$	$H = 10.34$, $p < 0.05^{*}$	$H = 11.70$, $p < 0.05^{*}$

the interface. This kind of exploration would entail a basic orientation of the space which would make the participant aware of the digit-buttons. The next two most common training classifications for PINs are just exploring the start point (Start Point Only, orienting to where to begin the PIN entry), or tracing out the entire gesture for entering all of the digits in the PIN (Traced Path).

For patterns, during training, unsurprisingly, the most common technique we observed was traces that retraced the path of the pattern. This makes sense given the swiping nature of the pattern. Partial exploration of the grid space and exploring just the start-point only are also quite common, but only half as common as retracing.

When considering the classifications during entry, the traces are divided into those that use the tactile aid and which do not. There was no difference statistically in the presence of a label (using χ^2) except for a significant decrease in Start-Hunt for patterns ($\chi^2 = 8.17$, $p < 0.005$). Since the tactile aid provided distinct feedback when the participant touched the start point, a decrease in the amount of hunting for the start point is expected. While there was a similar decrease for PIN entry using the aid, it was not significant ($\chi^2 = 1.31$, $p = 0.25$).

Table 10

Likert Responses to Ease of Tactile Aid: Response range from 1 to 5 (1, very easy, to 5, very hard).

	PIN	Pattern
Ease of learning	$\mu = 2.41$, $\sigma = 1.25$	$\mu = 1.89$, $\sigma = 0.93$
Ease of use	$\mu = 3.44$, $\sigma = 1.01$	$\mu = 2.74$, $\sigma = 1.13$

In Table 9, we analyze the impact of training label on performance. For PINs, the accuracy shows no significant differences across labels. However, we do see a difference for precision. Notably, both Start Point Only and Return to Start have much lower distances (higher precision). Additionally, those that performed a wide exploration had the highest distance (lower precision). This suggests that training with PINs should encourage a tighter training regime, perhaps confined to the start digit.

The training classification for patterns had significant differences for both accuracy and precision. Most notably, the Return to Start label showed much higher accuracy (0.91) and the lowest in the distance metric (highest precision). We also see strong performance gains for both Start Point Only and Traced Path. These results suggest that when designing training aids for patterns, just
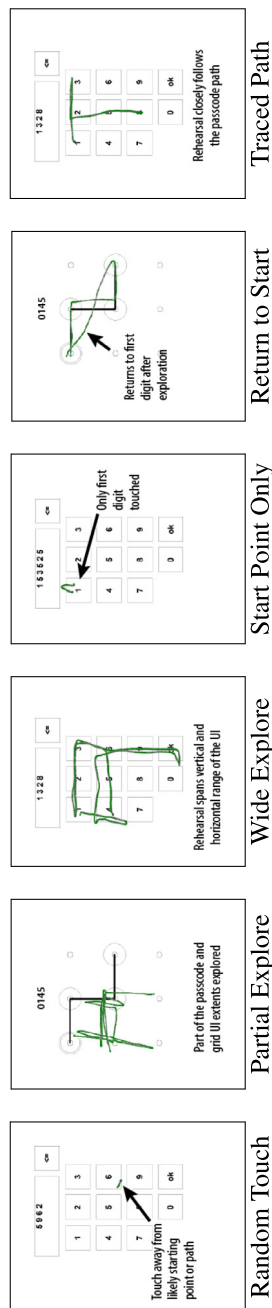
Fig. 4. Samples of training classification: see Table 8 for description of each classification type.

Fig. 5. Samples of entry classification: see Table 8 for description of each classification type.

like PINs, focusing on the start point is important. However, for patterns, encouraging a retrace of the path of the pattern can also be helpful.

4.5. Discussion from participants

Quantitative findings suggested that ease of learning and ease of use with the tactile aid were both harder for PIN compared to patterns (Table 10), In a similar vein, post-hoc discussion with participants revealed that challenges in the PIN condition were more frequently reported (n=7, three preferred the PIN aid operation). Several reasons for this were mentioned, most commonly the difficulty of accurately making long jumps between PIN numbers (e.g. 7 to 3 or 1 to OK, n=6). Participants also found more closely spaced PIN buttons a challenge (n=3), as well as accurately hitting the OK button at the end of a PIN sequence (n=3). Several specific passcode features, such as PIN digits located along the edge opposite a thumb for one-handed users (n=3), and consecutive numbers (n=1) were also mentioned as challenging factors.

More generally, participants reported the 6-digit PIN and grid conditions as being more difficult to perform (n=6), and felt the app itself might be hard to use discreetly in public (n=5). There was positive feedback on the utility of the distinct start point feature (n=4), which could guide the user to accurately begin their gesture. Regarding this feature, one participant stated that "it's all about the start" in maintaining spatialization throughout authentication gestures. Several participants also suggested giving the OK and Delete buttons a different vibrotactile coding to enhance the aid.

Two participants related feeling the aid worked best with a two-handed index pointer grip, but this challenged them because they typically gripped their phones one-handed, with a thumb pointer. Another participant, also typically a one- handed thumb grip user, noted that PIN jumps to central digits (e.g. 5 and 8) were much easier than jumps to digits on the outer edges (e.g. 1, 7, and 9).

5. Conclusion

This study was conducted to examine interaction techniques developed by users when they entered different types of passcodes on a mobile touchscreen device under eyes-free conditions (e.g., when worrying about the threat of observer attacks, or when facing situational impairments). We also inquired if tactile-only spatial feedback would effectively assist users with this type of screen unlocking. We did succeed, firstly, in capturing a picture of eyes-free authentication behaviors for common passcode entry methods. Looking at accuracy and precision measures, in particular, we can say regarding RQ1 that eyes-free unlocking overall (without tactile feedback) is understandably very challenging. PIN authentication is harder to perform with accuracy and precision, likely because of the numerous jumps the pointing finger must make. Looking at these measures for gestures that were aided by tactile feedback, as addressed by RQ2, we see a small positive effect on start point accuracy, for pattern unlocking. Otherwise, the tactile only feedback employed in this study is not helpful to eyes-free authentication in terms of accuracy, precision, or time taken to unlock. Movement traces were classified to better understand the strategies undertaken for unlocking and using tactile feedback, per RQ3. We found that strategies that focused on locating the first digit of the passcode were helpful. Precision for PIN entry was improved by the Start Point Only and Return to Start strategies. Similarly, accuracy and precision of pattern entry was helped by use of the Return to Start strategy. These classifications have been described, and can be used by other researchers for purposes of analyzing movement traces on mobile touchscreen technologies. These insights about user strategies, event and error types can also support the design of targeted authentication training aids for users who may frequently encounter similar eyes-free conditions.

Acknowledgments

### Appendix A. Patterns and PINs visualized

A1. Patterns

The double circle indicates a start point, single circles is a point included in the pattern. Note that labeling of patterns begins in the upper left with 0, incrementing across each row, ending in the lower right with 8. All visuals are also provided in `images/patterns` sub-directory.

| 0145 | 6745 | 3157 | 1346 | 4572 |
| 743521 | 136785 | 642580 | 014673 | 841257 |
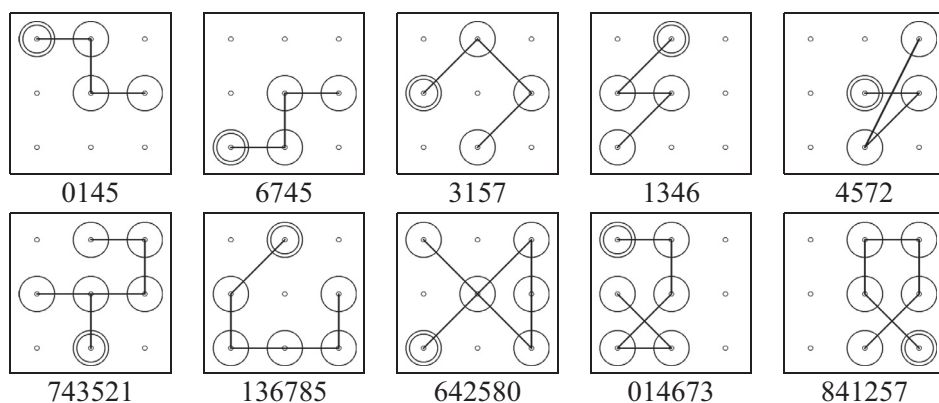
A2. PINs

The filled circle • indicates the start point, and unfilled circle ○ indicates an intermediate point. Line traces are provided to show expected shape and directionality of a trace, but users do not drag/maintain contact during entry. Rather, users enter the PIN as normally would be expected by clicking/pressing the buttons. All visuals are also provided in `images/pins` sub-directory.

| 1328 | 6702 | 1955 | 7272 | 5962 |
| 152525 | 458090 | 159428 | 441791 | 366792 |
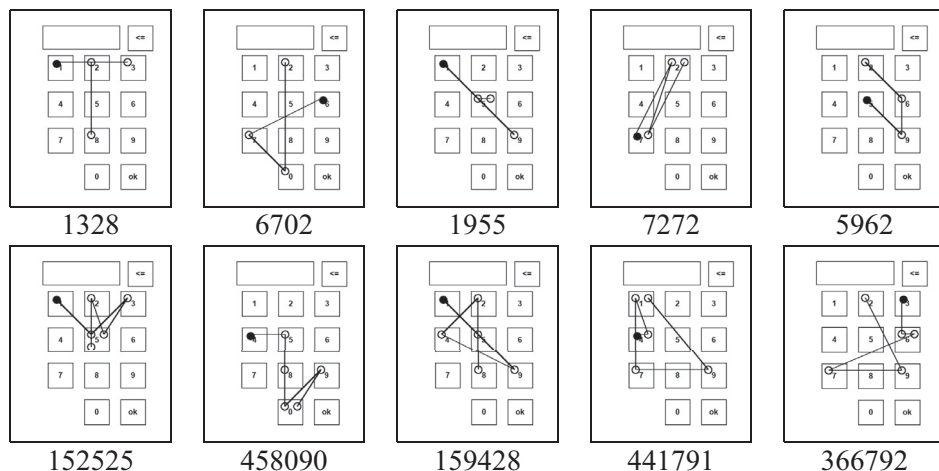
Appendix B. Pre-survey questions

- What is your age (18–24, 25–34, 35–44, 45–54, 55+, NA)?
- What is your identified gender?
- Do you have any physical conditions that might affect your ability to enter authentication passcodes on a mobile phone?
- Do you use a smartphone currently? What is its operating system? Why did you select that phone?
- Do you use an authentication method to lock your phone, and if so which method, and why (i.e. PIN, grid, TouchID, etc.)?
- Without telling me your current passcode, how do you select the passcodes you use to lock your phone (i.e. familiar number, or visual pattern)?
- How concerned are you with keeping your phone secure (1, not at all concerned, to 5, highly concerned)?
- What experiences can you recall involving people either trying to steal or use your phone without permission?
- What experiences can you recall involving people trying to observe your passcodes without permission?
- How concerned are you, typically, in a public space, with the threat of someone watching you authenticate and collecting your passcodes (1, not at all concerned, to 5, highly concerned)?
- If you had any of these experiences, how did it affect your behavior?
- Have any other experiences or concerns indirectly affected your authentication behavior (news articles, stories about friends, etc.)?
- If you do authentication, how do you typically hold your phone for that?

B1. Post-survey questions

- On a scale from 1 to 5, how difficult was entering passcode this way (1, very easy, to 5, very hard)? How so?
- On a scale from 1 to 5, how easy was the grid pattern tactile app to learn (1, very easy, to 5, very hard)? How so?
- On a scale from 1 to 5, how easy was the grid pattern tactile app to use (1, very easy, to 5, very hard)? How so?
- On a scale from 1 to 5, how easy was the PIN tactile app to learn (1, very easy, to 5, very hard)? How so?
- On a scale from 1 to 5, how easy was the PIN tactile app to use (1, very easy, to 5, very hard)? How so?
- Can you see yourself using the grid pattern tactile aid to help authenticate on your phone in your actual daily life? Why or why not?
- Can you see yourself using the PIN tactile aid to help authenticate on your phone in your actual daily life? Why or why not?
- How is this approach similar or different from how you enter passcodes on your phone now?
- Do you think the grid tactile aid would help protect you from someone shoulder surfing you? Why or why not?
- Do you think the PIN tactile aid would help protect you from someone shoulder surfing you? Why or why not?

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jisa.2018.05.004

References

[1] Abdolrahmani A, Kuber R, Hurst A. An empirical investigation of the situationally-induced impairments experienced by blind mobile device users. In: Proceedings of the thirteenth Web for all conference. New York, NY, USA: ACM; 2016 21:1-21:8. ISBN 978-1-4503-4138-7. http://doi.acm.org/10.1145/2899476.2899482

[2] Grinter RE, Palen L, Eldridge M. Chatting with teenagers: considering the place of chat technologies in teen life. ACM Trans Comput-Hum Interact 2006;13(4):423–47. doi:10.1145/1188816.1188817. http://doi.acm.org/10.1145/1188816.1188817

[3] Azenkot S, Bennett CL, Ladner RE. Digitaps: Eyes-free number entry on touch-screens with minimal audio feedback. In: Proceedings of the twenty-sixth annual ACM symposium on user interface software and technology. New York, NY, USA: ACM; 2013. p. 85–90. ISBN 978-1-4503-2268-3. doi:10.1145/2501988.2502056. http://doi.acm.org/10.1145/2501988.2502056

[4] Azenkot S, Lee NB. Exploring the use of speech input by blind people on mobile devices. In: Proceedings of the fifteenth international ACM SIGACCESS conference on computers and accessibility. New York, NY, USA: ACM; 2013 11:1-11:8. ISBN 978-1-4503-2405-2. http://doi.acm.org/10.1145/2513383.2513440

[5] Bonner MN, Brudvik JT, Abowd GD, Edwards WK. No-look notes: Accessible eyes-free multi-touch text entry. In: Pervasive, 10. Springer; 2010. p. 409–26. doi:10.1007/978-3-642-12654-3. https://doi.org/10.1007/978-3-642-12654-3

[6] Brewster S, Lumsden J, Bell M, Hall M, Tasker S. Multimodal 'eyes-free' interaction techniques for wearable devices. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2003. p. 473–80. ISBN 1-58113-630-7. doi:10.1145/642611.642694. http://doi.acm.org/10.1145/642611.642694

[7] Goel M, Findlater L, Wobbrock J. Walktype: Using accelerometer data to accomodate situational impairments in mobile touch screen text entry. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2012. p. 2687–96. ISBN 978-1-4503-1015-4. doi:10.1145/2207676.2208662. http://doi.acm.org/10.1145/2207676.2208662

[8] Gupta A., Samdaria N. Swift: Swift vision-free text-entry for touch screens. 2011. Technical report. Available: http://www.cs.utoronto.ca/~aakar/Publications/Svift-Report.pdf

[9] Li KA, Baudisch P, Hinckley K. Blindsight: Eyes-free access to mobile phones. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2008. p. 1389–98. ISBN 978-1-60558-011-1. doi:10.1145/1357054.1357273. http://doi.acm.org/10.1145/1357054.1357273

[10] Kajastila R, Lokki T. Eyes-free interaction with free-hand gestures and auditory menus. Int J Hum Comput Stud 2013;71(5):627–40. doi:10.1016/j.ijhcs.2012.11.003. https://doi.org/10.1016/j.ijhcs.2012.11.003

[11] Lu Y, Yu C, Yi X, Shi Y, Zhao S. Blindtype: eyes-free text entry on hand-held touchpad by leveraging thumb's muscle memory. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017;1(2) 18:1-18:24. doi:10.1145/3090083. http://doi.acm.org/10.1145/3090083

[12] Pielot M, Kazakova A, Hesselmann T, Heuten W, Boll S. Pocketmenu: Non-visual menus for touch screen devices. In: Proceedings of the fourteenth international conference on human-computer interaction with mobile devices and services. New York, NY, USA: ACM; 2012. p. 327–30. ISBN 978-1-4503-1105-2. doi:10.1145/2371574.2371624. http://doi.acm.org/10.1145/2371574.2371624

[13] Zhao S, Dragicevic P, Chignell M, Balakrishnan R, Baudisch P. Earpod: eyes-free menu selection using touch input and reactive audio feedback. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2007. p. 1395–404. ISBN 978-1-59593-593-9. doi:10.1145/1240624.1240836. http://doi.acm.org/10.1145/1240624.1240836

[14] Wolf F, Aviv AJ, Kuber R. Performance of eyes-free mobile authentication. In: Proceedings of the Workshop on Usable Security; 2018.

[15] von Zezschwitz E, Dunphy P, De Luca A. Patterns in the wild: a field study of the usability of pattern and pin-based authentication on mobile devices. In: Proceedings of the fifteenth international conference on human-computer interaction with mobile devices and services. New York, NY, USA: ACM; 2013. p. 261–70. ISBN 978-1-4503-2273-7. doi:10.1145/2493190.2493231. http://doi.acm.org/10.1145/2493190.2493231

[16] Yi B, Cao X, Fjeld M, Zhao S. Exploring user motivations for eyes-free interaction on mobile devices. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2012. p. 2789–92. ISBN 978-1-4503-1015-4. doi:10.1145/2207676.2208678. http://doi.acm.org/10.1145/2207676.2208678

[17] Ketabdar H, Moghadam P, Naderi B, Roshandel M. Magnetic signatures in air for mobile devices. In: Proceedings of the fourteenth international conference on human-computer interaction with mobile devices and services companion. New York, NY, USA: ACM; 2012. p. 185–8. ISBN 978-1-4503-1443-5. doi:10.1145/2371664.2371705. http://doi.acm.org/10.1145/2371664.2371705

[18] Oakley I, Park J-S. Designing eyes-free interaction. Haptic and audio interaction design 2007:121–32. doi:10.1007/978-3-540-76702-2.

[19] Tinwala H, MacKenzie IS. Eyes-free text entry with error correction on touch-screen mobile devices. In: Proceedings of the sixth nordic conference on human-computer interaction: extending boundaries. New York, NY, USA: ACM; 2010. p. 511–20. ISBN 978-1-60558-934-3. doi:10.1145/1868914.1868972. http://doi.acm.org/10.1145/1868914.1868972

[20] Azenkot S, Rector K, Ladner R, Wobbrock J. Passchords: Secure multi-touch authentication for blind people. In: Proceedings of the fourteenth international acm sigaccess conference on computers and accessibility. New York, NY, USA: ACM; 2012. p. 159–66. ISBN 978-1-4503-1321-6. doi:10.1145/2384916.2384945. http://doi.acm.org/10.1145/2384916.2384945

[21] Guerreiro T, Jorge J, Gonçalves D. Exploring the non-visual acquisition of targets on touch phones and tablets. In: Proceedings of the second workshop on mobile accessibility, MobileHCI; 2012.

[22] Guerreiro T, Montague K, Guerreiro Ja, Nunes R, Nicolau H, Gonçalves DJ. Blind people interacting with large touch surfaces: Strategies for one-handed and two-handed exploration. In: Proceedings of the 2015 international conference on interactive tabletops & surfaces. New York, NY, USA: ACM; 2015. p. 25–

34. ISBN 978-1-4503-3899-8. doi:10.1145/2817721.2817743. http://doi.acm.org/10.1145/2817721.2817743

[23] Kane SK, Morris MR, Wobbrock JO. Touchplates: Low-cost tactile overlays for visually impaired touch screen users. In: Proceedings of the fifteenth international ACM SIGACCESS conference on computers and accessibility. New York, NY, USA: ACM; 2013 22:1-22:8. ISBN 978-1-4503-2405-2. http://doi.acm.org/10.1145/2513383.2513442

[24] McGookin D, Brewster S, Jiang W. Investigating touchscreen accessibility for people with visual impairments. In: Proceedings of the fifth nordic conference on human-computer interaction: building bridges. New York, NY, USA: ACM; 2008. p. 298–307. ISBN 978-1-59593-704-9. doi:10.1145/1463160.1463193. http://doi.acm.org/10.1145/1463160.1463193

[25] McGookin D, Brewster S, Jiang W. Investigating touchscreen accessibility for people with visual impairments. In: Proceedings of the fifth nordic conference on human-computer interaction: building bridges. New York, NY, USA: ACM; 2008. p. 298–307. ISBN 978-1-59593-704-9. doi:10.1145/1463160.1463193. http://doi.acm.org/10.1145/1463160.1463193

[26] De Luca A, Hang A, Brudy F, Lindner C, Hussmann H. Touch me once and i know it's you!: Implicit authentication based on touch screen patterns. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2012. p. 987–96. ISBN 978-1-4503-1015-4. doi:10.1145/2207676.2208544. http://doi.acm.org/10.1145/2207676.2208544

[27] Bianchi A, Oakley I, Lee JK, Kwon DS. The haptic wheel: Design & evaluation of a tactile password system. In: Proceedings of the extended abstracts on human factors in computing systems. New York, NY, USA: ACM; 2010. p. 3625–30. ISBN 978-1-60558-930-5. doi:10.1145/1753846.1754029. http://doi.acm.org/10.1145/1753846.1754029

[28] De Luca A, von Zezschwitz E, Hussmann H. Vibrapass: secure authentication based on shared lies. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: ACM; 2009. p. 913–16. ISBN 978-1-60558-246-7. doi:10.1145/1518701.1518840. http://doi.acm.org/10.1145/1518701.1518840

[29] Ali A, Aviv AJ, Kuber R. Developing and evaluating a gestural and tactile mobile interface to support user authentication. IConference 2016 Proceedings 2016. doi:10.9776/16141. https://doi.org/10.9776/16141

[30] Tinwala H, MacKenzie IS. Letterscroll: text entry using a wheel for visually impaired users. In: Proceedings of the extended abstracts on human factors in computing systems. New York, NY, USA: ACM; 2008. p. 3153–8. ISBN 978-1-60558-012-8. doi:10.1145/1358628.1358823. http://doi.acm.org/10.1145/1358628.1358823

[31] von Zezschwitz E, De Luca A, Janssen P, Hussmann H. Easy to draw, but hard to trace?: On the observability of grid-based (un)lock patterns. In: Proceedings of the thirty-third annual ACM conference on human factors in computing systems. New York, NY, USA: ACM; 2015. p. 2339–42. ISBN 978-1-4503-3145-6. doi:10.1145/2702123.2702202. http://doi.acm.org/10.1145/2702123.2702202

[32] Aviv AJ, Davin JT, Wolf F, Kuber R. Towards baselines for shoulder surfing on mobile authentication. In: Proceedings of the thirty-third annual computer security application conference. New York, NY, USA: ACM; 2017.

[33] Davin JT, Aviv AJ, Wolf F, Kuber R. Baseline measurements of shoulder surfing analysis and comparability for smartphone unlock authentication. In: Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems. New York, NY, USA: ACM; 2017. p. 2496–503. ISBN 978-1-4503-4656-6. doi:10.1145/3027063.3053221. http://doi.acm.org/10.1145/3027063.3053221

[34] Bonneau J, Preibusch S, Anderson R. A birthday present every eleven wallets? the security of customer-chosen banking PINs. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 25–40. https://doi.org/10.1007/978-3-642-32946-3_3

[35] Aviv AJ, Budzitowski D, Kuber R. Is bigger better? comparing user-generated passwords on 3x3 vs. 4x4 grid sizes for android's pattern unlock. In: Proceedings of the thirty-first annual computer security applications conference. New York, NY, USA: ACM; 2015. p. 301–10. ISBN 978-1-4503-3682-6. doi:10.1145/2818000.2818014. http://doi.acm.org/10.1145/2818000.2818014

[36] Yang Y, Clark GD, Lindqvist J, Oulasvirta A. Free-form gesture authentication in the wild. In: Proceedings of the 2016 CHI conference on human factors in computing systems. New York, NY, USA: ACM; 2016. p. 3722–35. ISBN 978-1-4503-3362-7. doi:10.1145/2858036.2858270. http://doi.acm.org/10.1145/2858036.2858270

[37] Liu C, Clark GD, Lindqvist J. Guessing attacks on user-generated gesture passwords. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017;1(1) 31-3:24. doi:10.1145/3053331. http://doi.acm.org/10.1145/3053331

[38] Clark GD, Lindqvist J. Engineering gesture-based authentication systems. IEEE Pervasive Comput 2015;14(1):18–25. doi:10.1109/MPRV.2015.6.

[39] Micinski K, Votipka D, Stevens R, Kofinas N, Mazurek ML, Foster JS. User interactions and permission use on android. In: Proceedings of the 2017 CHI conference on human factors in computing systems. New York, NY, USA: ACM; 2017. p. 362–73. ISBN 978-1-4503-4655-9. doi:10.1145/3025453.3025706. http://doi.acm.org/10.1145/3025453.3025706