Using a Blocklist to Improve the Security of User Selection of Android Patterns

Collins W. Munyendo*, Miles Grant*, Philipp Markert‡, Timothy J. Forman§, and Adam J. Aviv*
*The George Washington University, ‡Ruhr University Bochum, §United States Navy

Abstract

Android patterns remain a popular method for unlocking smartphones, despite evidence suggesting that many users choose easily guessable patterns. In this paper, we explore the usage of blocklists to improve the security of user-chosen patterns by disallowing common patterns, a feature currently unavailable on Android but used by Apple during PIN selection. In a user study run on participants' smartphones ($n = 1006$), we tested 5 different blocklist sizes and compared them to a control treatment. We find that even the smallest blocklist (12 patterns) had benefits, reducing a simulated attacker's success rate after 30 guesses from 24 % to 20 %. The largest blocklist (581 patterns) reduced the percentage of correctly guessed patterns after 30 attempts down to only 2 %. In terms of usability, blocklists had limited negative impact on short-term recall rates and entry times, with reported SUS values indicating reasonable usability when selecting patterns in the presence of a blocklist. Based on our simulated attacker performance results for different blocklist sizes, we recommend blocking 100 patterns for a good balance between usability and security.

1 Introduction

Restricting access to smartphones is critical for security, as these devices play an important role in our daily lives. A common method to secure smartphone access is unlock authentication, such as using a PIN or password, that the user enters to unlock the device. On Android devices, users can also choose to select a graphical method in the form of unlock patterns, where users enter a previously selected pattern by swiping on a 3x3 grid of points.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2021.
August 8–10, 2021, Virtual Conference.

While user-selected passwords and PINs for mobile authentication have been shown to be more resilient to guessing attacks compared to Android patterns [19], patterns remain popular among a large group of Android users. Our study finds that about 27 % of participants use patterns, matching inquiries from prior work [4, 14, 17, 19]. Even though 3x3 patterns allow for 389,112 options, more than the 10,000 choices offered by 4-digit PINs, users select from a much smaller subset of patterns that are easily predicted and would be guessed by an informed attacker, even after just 30 attempts [4, 28].

There have been several proposals in the past to improve the security of Android patterns. Some suggest using ad-hoc strength meters [2, 24, 25], feedback during selection [12], rearrangement of the contact points [27, 28], or expansion from a 3x3 to a 4x4 grid [4]. However, these suggestions all have their drawbacks. For instance, increasing the grid size has proven not to increase security significantly for an online attacker with a few guesses [4]. Other proposals, including the rearrangement of the grid, change the simple interface that makes Android patterns so popular in the first place.

To address these challenges, we propose using blocklists during pattern selection. This feature, which is used by Apple iOS devices during PIN selection, disallows common options so that users select more diversely. Recent research on mobile authentication PINs [19] and Knock Codes [21] indicates that a well sized blocklist can have significant improvements on security with limited impact on usability. In this paper, we ask *(a) what is the security and usability impact of blocklists on Android unlock patterns, and (b) what is the "right" sized blocklist that balances security and usability?*

To answer these questions, we carried out an online survey on Amazon Mechanical Turk with $n = 1006$ participants. Participants were assigned to 1 of 6 treatments, a control as well as 5 blocklist-enforcing treatments. For the latter, we varied the size of the blocklists with patterns chosen based on prior studies [4, 18, 28, 30]. In the course of the survey, participants created and recalled a pattern, answered questions about the general usability and described their strategies for selecting their pattern. Participants that encountered a blocklist were

additionally asked about changes to their selection strategy while those that did not were asked if their strategy would change upon encountering a blocklist warning.

We evaluated the security of unlock patterns selected across different treatments using guessability metrics. We primarily considered a throttled attacker scenario as it is the most relevant for mobile authentication where an attacker only has 10–30 guesses before a lockout or at least significant delays ($>$ 1 hr) occur on the device. We find that that 24 % of patterns in the control treatment are guessed after 30 guesses. In contrast, the smallest blocklist reduces the attacker's performance to 20 %, and with the largest blocklist in place, the attacker only guesses 2 % of the patterns within 30 attempts.

For usability, blocklists had minimal impact on short-term recall rates. While the average selection time increases due to the interaction with the blocklist (a one time cost), changes in entry times are negligible. Participants in the largest blocklist treatments only took, on average, an additional 0.26 seconds to enter their patterns. Participant responses evaluated using the System Usability Scale (SUS) support these findings, with scores ranging from 78.6 for the control to 71.6 for the largest blocklist treatment, indicating that the addition of blocklists improves security while appearing not to have meaningful effects on the usability of unlock patterns. However, additional work is required to explore long-term recall rates.

To summarize, we make the following contributions:

1. We study the effects of blocklists on Android unlock patterns, showing that they are able to significantly increase security even for small blocklist sizes. Patterns selected in the blocklist treatments are harder to guess for both a simulated and a perfect knowledge attacker.
2. We show that blocklists might not have meaningful effects on the usability of unlock patterns. The SUS scores, entry times, and short-term recall rates across all 5 blocklist treatments are comparable to the control treatment.
3. We provide guidance to improve the existing implementation of unlock patterns, with our results suggesting that a blocklist containing the 100 most common patterns improves the security of user-chosen patterns while appearing to minimally impact their short-term recall rates.

2 Related Work

Android unlock patterns, first introduced in 2008 as a modification of the Pass-Go scheme [26], are one of the most widely used knowledge-based authentication mechanisms on smartphones today. Despite being less secure than PINs [1, 4, 10, 19, 28, 31] or passwords [9, 22], 27 % of participants in our study use patterns to secure their smartphones, which matches inquiries from prior work [4, 14, 17, 19].

Some of the security limitations of patterns were first demonstrated by Uellenbeck et al. [28]. Through a large scale user study measuring users' actual choices of patterns, Uellenbeck et al. found that selection strategies for patterns are biased, including a preference to start from the top left corner and end in the bottom right corner of the grid. Loge et al. [18] found that personal traits of a user influence the strength of unlock patterns they select. Other studies have shown patterns to be vulnerable to smudge attacks [6, 11], shoulder surfing attacks [5, 8, 22], sensor attacks [7], video attacks [32], and physical attacks [3].

As a workaround, there have been several proposals to improve the security of Android unlock patterns. Some of these suggestions include the use of strength meters during pattern selection [2, 24, 25], rearrangement of the grid points [27, 28], use of background images during pattern selection [30], modification of the pattern size to prevent various attacks [13,23,29] , forcing users to choose certain points during pattern selection [12], or the use of Double Patterns [14]. However, all these suggestions have their drawbacks. For instance, increasing the grid size has been shown not to improve security [4] and strength meters are constrained by the inaccuracy of their underlying algorithms [15]. It is also unclear if methods that fundamentally change pattern entry, like Double Patterns [14] or Pass-O [27], will have widespread user support or adoption, despite security benefits.

Here, we propose using blocklists, which do not change the input interface, and have evidence of positive security effects, such as by Markert et al. [19] for PINs, Samuel et al. [21] for Knock Codes, and Forman and Aviv [14] for Double Patterns. Markert et al. [21] found that a small, enforcing blocklist would have large effects on PIN guessability, and that a blocklist of approximately 1000 PINs would properly balance usability and security. Forman and Aviv [14] found that small blocklists of first-pattern selection for Double Patterns had a similarly outsized effect on security, and Samuel et al. [21] found that blocklists significantly improve the security of Knock Codes. While our study similarly explores the security and usability of blocklists on smartphone authentication, it differs from the above studies by focusing on traditional, unmodified Android unlock patterns. In the end, we find that blocklists, even relatively small ones, can significantly improve the security of unlock patterns, inline with prior results [14, 19, 21].

3 Methodology

In the following, we describe our methodology. We start by outlining the design of the user study and giving a detailed description of the 6 treatments, and following, we discuss the recruitment process, limitations, and ethics.

3.1 Survey Structure

We conducted an online survey on Amazon Mechanical Turk (MTurk), and to ensure ecological validity of selecting and entering patterns on a mobile device, the survey was designed

to be taken on mobile browsers only, as checked via the user-agent. Our study was open to both pattern and non-pattern users, with pattern users free to select patterns they already use unless blocked as part of a blocklist treatment. Participants were assigned 1 of 6 treatments for selecting and recalling a pattern: 5 blocklist-enforcing treatments with blocklists of various sizes and 1 control treatment without any blocklist intervention. We will discuss those treatments in more detail in Section 3.2. The average time to complete the survey was 6 minutes and participants were compensated $1.00.

We will now outline the structure of the survey; for a detailed description, please refer to Section A in the Appendix.

1. *Informed Consent*: Participants were informed about the purpose, structure, and anticipated duration of the study as well as the compensation.
2. *Device Usage*: Participants were asked about the number of smartphones they use, the device brands, and their authentication methods. Details regarding device usage can be found in Table 7.
3. *Background Information*: Because we could not expect all participants to be familiar with Android patterns, we provided background information including how to create a valid pattern. We further showed them an image with the Android unlock pattern interface, but not an entered pattern to avoid priming.
4. *Practice*: Participants were asked to practice creating a pattern before proceeding, serving as a hands-on introduction to patterns. The patterns selected here are not used in our analysis.
5. *Instructions/Scenario-Priming*: After familiarizing with Android unlock patterns, participants were informed that they should now create a pattern that they would use to secure their primary smartphone. Participants were also informed that they would have to recall the pattern they selected and therefore, it would need to be both secure and memorable. Participants were additionally instructed not to write down or use any aids to help them remember their pattern. To proceed, participants had to confirm that they understood all of the mentioned instructions.
6. *Selection/Blocklist-Intervention*: Participants selected (and confirmed) a pattern as they would use to secure their primary smartphone. During selection, participants in the blocklist treatments saw the warning depicted in Figure 1 if they entered a disallowed pattern and were asked to select a different pattern.
7. *SUS*: After selecting a pattern, participants answered questions from the System Usability Scale (SUS) to determine their perceived usability of pattern selection.
8. *Post-Entry*: In addition to the SUS questions, participants were asked whether they felt they created a pattern that provides adequate security and whether it was difficult for them to select the pattern.
9. *Strategy*: To understand how users select and change their patterns, we asked participants that encountered a blocklist for their selection strategy prior to the warning and how their strategy changed after seeing it. Participants that did not encounter a blocklist were asked to imagine how their strategy would change if they encountered the blocklist warning.
10. *Recall*: Participants attempted to recall their pattern within 5 attempts.
11. *Security Comparison*: After recall, participants were asked about the security of patterns in general and in comparison to 3 other unlock methods: 4-digit, as well as 6-digit PINs, and alphanumeric passwords.
12. *Real World Usage*: To better understand whether the patterns created in the study would actually be used, we asked participants if they would select the same unlock pattern on their smartphones along with their reasons for that decision.
13. *Demographics*: Participants were asked to provide demographic information, such as age, identified gender, dominate hand, education, and technical background. We also included a second attention check question on this page. To ensure that the demographic backgrounds of the participants do not interfere with the rest of the study [20], we asked these questions at the very end of the study.
14. *Honesty*: Finally, we asked participants if they had honestly participated in the survey and followed instructions completely. We paid all participants who completed the study but discarded participants from the analysis if they indicated dishonesty at this point.

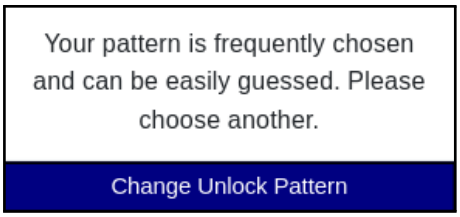


Figure 1: Blocklist warning used in the study.

3.2 Treatments

Participants were randomly assigned to either a control treatment or 1 of the 5 blocklist-enforcing treatments. To determine the common patterns to block, we combined data from von Zezschwitz et al. [30], Aviv et al. [4], Uellenbeck et al. [28], and Loge et al. [18], for a total of 4,637 patterns. Blocklists were generated by selecting patterns that appeared at least a certain number of times in the data set, e.g., at least 2 times for BL-2 (the largest blocklist with 581 patterns) or at least 32 times BL-32 (the smallest blocklist with 12 patterns).[1] The treatments are described below:

[1]To foster future research on this topic, we share the described blocklists. Please contact the authors for this purpose.

- **Control** ($n = 169$): Participants received no interventions when selecting a pattern.
- **BL-2** ($n = 166$)**:** The blocklist in this treatment comprised 581 patterns, with these patterns appearing at least twice in prior work.
- **BL-4** ($n = 172$)**:** The blocklist in this treatment comprised 239 patterns, with these patterns appearing at least 4 times in prior work.
- **BL-8** ($n = 161$)**:** The blocklist in this treatment comprised 105 patterns, with these patterns appearing at least 8 times in prior work.
- **BL-16** ($n = 165$)**:** The blocklist in this treatment comprised 54 patterns, with these patterns appearing at least 16 times in prior work.
- **BL-32** ($n = 173$)**:** The blocklist in this treatment comprised 12 patterns, with these patterns appearing at least 32 times in prior work.

Participants in the blocklist treatments received the warning message in Figure 1 when they selected a blocked pattern, which is based on the iOS blocklist warning [14, 19]. Blocklists were enforcing, i.e., could not be ignored, and participants were required to select a pattern that was not blocked.

3.3 Recruitment and Demographics

We recruited $n = 1006$ participants on Amazon Mechanical Turk (MTurk), after excluding 65 responses due to failed attention checks or dishonesty. As expected when recruiting from MTurk, our surveyed population was comprised primarily of younger (59 % between 18–34), male-identifying (62 % male, 36 % female, and 2 % other gender, or prefer not to say) participants with semi- or full college education (28 % some college or Associate's, 60 % Bachelor's or above). Table 1 depicts the full demographic information.

Table 1: Demographic information of participants.

	Male		Female		Other		Total	
	No.	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Age	624	62 %	367	36 %	15	1 %	1006	100 %
18–24	80	8 %	46	5 %	1	0 %	127	13 %
25–29	150	15 %	84	8 %	3	0 %	237	24 %
30–34	137	14 %	79	8 %	1	0 %	217	22 %
35–39	106	11 %	69	7 %	4	0 %	179	18 %
40–44	54	5 %	25	2 %	0	0 %	79	8 %
45–49	48	5 %	27	3 %	0	0 %	75	7 %
50–54	27	3 %	11	1 %	0	0 %	38	4 %
55–59	9	1 %	11	1 %	0	0 %	20	2 %
60–64	5	0 %	6	1 %	0	0 %	11	1 %
65+	8	1 %	9	1 %	0	0 %	17	2 %
Prefer not to say	0	0 %	0	0 %	6	1 %	6	1 %
Education	624	62 %	367	36 %	15	1 %	1006	100 %
Some High Sch.	0	0 %	2	0 %	0	0 %	2	0 %
High School	56	6 %	29	3 %	0	0 %	85	8 %
Some College	119	12 %	66	7 %	1	0 %	186	18 %
Trade	17	2 %	9	1 %	0	0 %	26	3 %
Associate's	51	5 %	44	4 %	1	0 %	96	10 %
Bachelor's	288	29 %	168	17 %	5	0 %	461	46 %
Master's	74	7 %	41	4 %	1	0 %	116	12 %
Professional	10	1 %	5	0 %	0	0 %	15	1 %
Doctorate	9	1 %	3	0 %	0	0 %	12	1 %
Prefer not to say	0	0 %	0	0 %	7	1 %	7	1 %
Background	624	62 %	367	36 %	15	1 %	1006	100 %
Technical	266	26 %	87	9 %	1	0 %	354	35 %
Non-Technical	335	33 %	265	26 %	4	0 %	604	60 %
Prefer not to say	23	2 %	15	1 %	10	1 %	48	5 %

3.4 Limitations

Our study has a number of limitations. Due to the nature of online surveys, it is not possible to tell whether participants fully and completely followed the instructions provided in the survey. We tried to mitigate this by including 2 attention check questions in the survey and asking participants whether or not they answered honestly, highlighting that they would be paid irrespective of their answer. Additionally, we reviewed all participant responses and removed participants from our analysis whose responses were inconsistent. As with other studies, participants on MTurk tended to be younger and more educated. We do not make any claims about our results being representative of the general population. As our study was relatively short, the recall rates reflect short-term memorability of unlock patterns; future work is needed to explore long-term memorability of these patterns. However, this approach has been used with a lot of success by many other researchers in the community to study mobile authentication [3, 4, 14, 18, 19, 28].

This survey may have been participants' first exposure to unlock patterns (27 % of participants were pattern users), and as a result, the non-pattern users' selection may vary in a real-world setting. To test for this, we asked non-pattern users if they would use the pattern they created to secure their primary smartphone. The results show that 40 % would use the pattern they created, 26 % were unsure and 34 % would not. Most participants who indicated that they were unsure or would not use their pattern argued that they would use it, had it not been recorded in the survey. This suggests that the patterns of the participants who have not previously used this unlock method closely match up to pattern selection in the real world.

3.5 Ethical Considerations

The study was approved by our Institution's Review Board (IRB), and participants were fully informed about the purpose and structure of the study. All participants were paid regard-

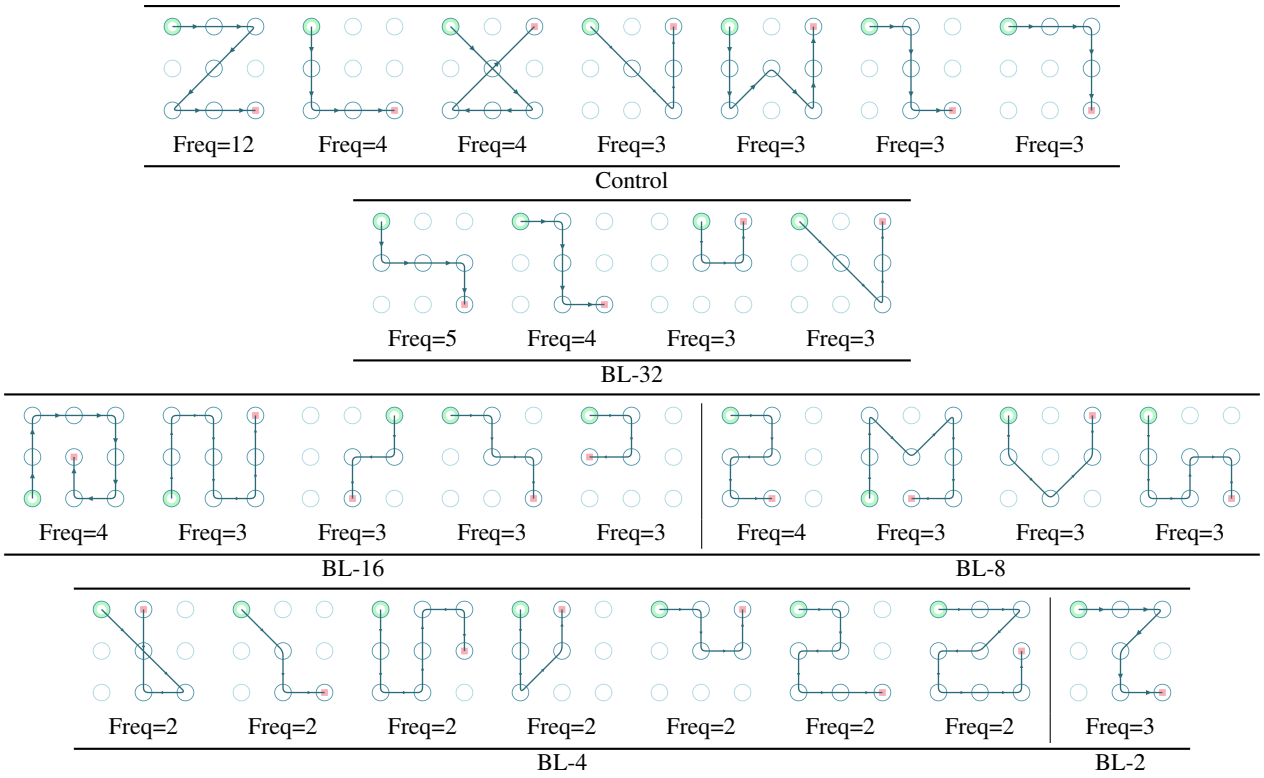


Figure 2: Most frequent patterns observed in treatments. Green circles depict start of a pattern, red squares indicate the end.

less of the quality of their submitted data. This includes cases where we removed submissions from our analysis for failing attention checks as well as situations where participants indicated dishonesty in answering the survey questions.

Another aspect to consider is the risk associated with the login information participants share with us through the study. As described earlier, some participants said they would use the unlock patterns they created in the study and others even confirmed that they do use the very same pattern to secure their smartphone. While a targeted attacker could potentially use this information to harm users, the implied risk is minimal. There is no identifiable connection from the selected unlock patterns to individual participants. On the other hand, this research offers much benefit as the outcomes of improved blocklists assist future selection of Android unlock patterns.

4 Features of Collected Patterns

In this section, we discuss pattern properties including the most frequent patterns, lengths of the patterns as well as common start and end points of patterns selected by participants across treatments. Figure 2 shows the most frequent patterns selected across the different treatments. The start point of each pattern is a green circle while the end point is a red square. An arrow indicates the direction of the pattern from the start point to the end point.

The most common patterns in the control treatment directly depict letters such as Z ($n = 12$), L ($n = 4$), and W ($n = 3$), along with patterns that resemble letters such as X ($n = 4$) or V ($n = 3$). The latter is also popular in the BL-32 treatment along with a small U ($n = 3$), but the most popular are 2 patterns which start in the upper left, move through the central point, and end in the lower right. In BL-16, flipped letters including G ($n = 4$), U ($n = 3$) and the number 2 ($n = 3$) are more common. The number 2 ($n = 5$) is the most frequent pattern in BL-8, followed by the letter V ($n = 3$). We also observe modifications to letters in this treatment, including addition of lines to the letter M ($n = 3$). Patterns in the BL-4 treatment are notably more diverse, with the common patterns only appearing twice at most, including more advanced modification of letters such as U ($n = 2$), Z ($n = 2$), and number 2 ($n = 2$). Similarly, patterns in BL-2 are more diverse, with a more advanced modification to letter Z ($n = 2$) being the only pattern appearing at least twice.

Many patterns in the control treatment appear to use shapes including numbers or letters in their exact form factor, while shapes are altered by flipping, mirroring or adding extra lines in the blocklist treatments. This is further confirmed through our qualitative analysis of users' pattern selection strategies: most participants initially select their patterns based on shapes such an initial of their name for memorability, but add complexity, such as extra lines, when they encounter a blocklist.

Table 2: Properties of selected patterns.

Treatment	Patterns No.	Unique Patterns No.	Unique Patterns %	Blocklist Hits No.	Blocklist Hits Average	Participants with Hits No.	Participants with Hits %	Length Average	Length Std. Dev.	Stroke Length Average	Stroke Length Std. Dev.
Control	169	130	77 %	0	0.00	0	0.0 %	6.1	1.5	5.4	1.7
BL-32	166	142	86 %	41	0.25	32	19.3 %	6.1	1.5	5.4	1.7
BL-16	172	151	88 %	98	0.57	61	35.5 %	6.1	1.7	5.4	1.8
BL-8	161	141	88 %	129	0.80	66	41.0 %	6.0	1.6	5.4	1.7
BL-4	165	158	96 %	202	1.22	86	52.1 %	6.3	1.7	5.7	1.9
BL-2	173	155	90 %	368	2.13	127	73.4 %	6.2	1.5	5.4	1.6
Total	1006	724	72 %	838	0.83	372	37.0 %	6.1	1.6	5.5	1.7

Figure 3: Frequency of start and end points. The top row shows the start points while the bottom shows the end points.

Figure 3 shows the most common start and end points for patterns, with the top row depicting start points and the bottom row depicting end points. As reported in prior work [3, 4, 28], most patterns start in the top left corner of the grid and end in the bottom right. However, this becomes less prevalent for patterns selected in blocklist treatments, with 36.0 % to 44.6 % of these patterns starting in the top left corner, instead of 49.1 % in the control treatment. For the end points, 15.7 % to 22.3 % of patterns in the blocklist treatments end in the bottom right corner compared to 32.5 % in the control group. While there was no significant difference in starting at the top left corner, a *chi*-square test showed significant difference in ending at the bottom right corner ($\chi = 17.65$, $p < 0.01$) across treatments. This suggests that blocklists likely pushed participants to change their end points more as compared to their start points when encountering a blocklist.

We also considered the lengths of patterns, both in terms of number of contact points used (i.e., *length*) and the length of the strokes within the pattern (i.e., *stroke length*) (see Table 2). The stroke length is calculated by taking the Cartesian difference with the origin mapped to the center point and unit distances between points. We find no significant differences for length ($f = 0.639, p = 0.66$) or stroke-length ($f = 0.937, p = 0.45$) between treatments. Participants select patterns of similar lengths, but varied other properties after encountering blocklists.

Finally, we compared the number of unique patterns across different treatments. As shown in Table 2, the percentage of unique patterns selected by participants increases with the blocklist size. While only 77 % of the patterns were unique in the control treatment, 90 % were in BL-2, the largest blocklist size, and even 96 % in the second largest blocklist BL-4. We performed a χ^2 test on the prevalence of unique patterns, finding there is a significant difference ($\chi = 11.04$, $p = 0.05$). Post-hoc analysis (Bonferoni-corrected) revealed that only BL-2 and Control were significantly different, where BL-2 had the highest rate of unique patterns.

5 Security Analysis

In this section, we describe the security analysis of patterns selected with and without a blocklist. First, we introduce the the attacker model for a perfect knowledge and simulated attacker, and then we discuss the success rates of the 2 guessing attacks. Lastly, we discuss an analysis of selecting a blocklist size that balances the security and usability of patterns.

Attacker Model. We make a number of assumptions for our attacker model. Foremost, the attacker is generic and does not have additional information about individual users to perform a targeted attack. Such an attacker could use tailored techniques, for example, shoulder surfing [5, 8, 22] or smudge attacks [6], which may increase the success rate for a given victim, but be less successful in general.

We also consider 2 variations of the generic attacker, a perfect knowledge and a simulated attacker. The perfect knowledge attacker provides an upper bound performance of the generic attacker since it assumes the attacker knows the exact distribution of frequencies of patterns, and always guesses the next most frequent pattern. On the other hand, a simulated attacker utilizes a set of training data to guess an unknown set of the authentication. For the simulated attacker, we also assume that the attacker has knowledge of the blocklist and optimizes the guessing order by skipping patterns which could not be selected. This is because an attacker would have access to the best training material, including the blocked patterns. For the perfect knowledge attacker, this assumption is always implied as the attacker is aware of the distribution.

5.1 Perfect Knowledge Attacker

The perfect knowledge attacker results are presented in Table 3. To control the different sizes of our treatment groups and allow for a fair comparison, we randomly down-sampled all larger data sets to 161, i.e., the size of the BL-8 treatment. For the strength estimations of the perfect knowledge attacker, we use two metrics, β-*success-rate* and α-*guesswork*, as defined by Bonneau [9].

First, for an attacker that is limited in the number of guesses as is the case with unlock patterns, β-success-rate describes the percentage of the dataset guessed after β guesses. Reported as λ_β in Table 3, it is evident that blocklists greatly reduce the success rate of such a throttled attacker. BL-4, the second largest blocklist size, appears to reduce the attacker performance the most across the scenarios we investigated. After 3 guesses, BL-32 (the smallest blocklist) reduces the attacker performance from 13.1 % down to 9.0 % of patterns successfully guessed, as compared to the control treatment. BL-4 reduces the attacker performance even further, with only 4.6 % of patterns guessed after 3 attempts. After 10 guesses, BL-32 reduces the attacker performance from 22.9 % to 18.0 % compared to the control group. BL-4 further reduces the attacker success rate to only 10.9 % of patterns guessed after 10 attempts. After 30 guesses, the attacker can guess 41.1 % of patterns in the control group, but only 33.1 % in BL-32 and 22.3 % in BL-4. This suggests that even small blocklists can improve the security of user-selected patterns.

The second metric, α-guesswork, measures an attacker who is not constrained by the number of attempts to guess an authentication, otherwise known as an *unthrottled attacker*. Using bits of entropy, it measures how much "work" is required to guess an α fraction of the data set. A higher entropy implies more work for the attacker and ultimately shows the authentication is stronger. These results are indicated by $\widetilde{G}_\alpha$ in Table 3. Across all cases, the attacker is less successful when guessing patterns in the blocklist treatments compared to the control group. Just like in the throttled setting, patterns selected in the BL-4 treatment are stronger, with the α-guesswork being higher compared to the other groups for all guessing scenarios evaluated. When guessing 50 % of the data, BL-32, the smallest blocklist increases the guessing entropy by 0.33 compared to the control treatment. BL-4 further increases the entropy by 0.68 as compared to control. This again advocates that blocklists increase security of unlock patterns.

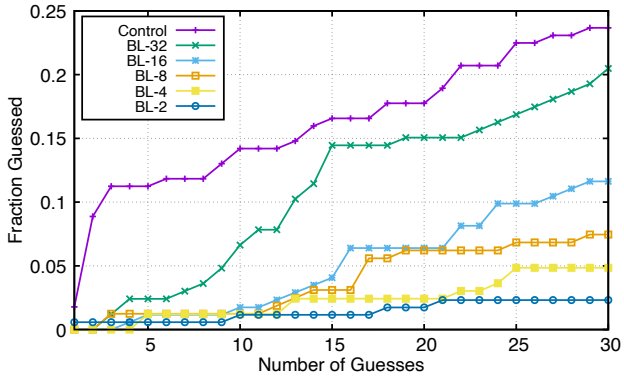
Table 3: Guessing metrics for a perfect knowledge attacker.

	Throttled Attack (%)			Unthrottled Attack (Bits)			
Treatment	λ_3	λ_{10}	λ_{30}	H_∞	$\widetilde{G}_{0.1}$	$\widetilde{G}_{0.3}$	$\widetilde{G}_{0.5}$
Control	13.1 %	22.9 %	41.1 %	3.75	4.66	6.00	6.93
BL-32	9.0 %	18.0 %	33.1 %	5.01	5.82	6.65	7.26
BL-16	7.3 %	15.6 %	29.9 %	5.33	6.04	7.00	7.45
BL-8	8.0 %	17.1 %	31.9 %	5.33	5.89	6.81	7.33
BL-4	4.6 %	10.9 %	22.3 %	6.33	6.64	7.34	7.61
BL-2	5.1 %	13.1 %	27.9 %	5.75	6.31	7.00	7.43

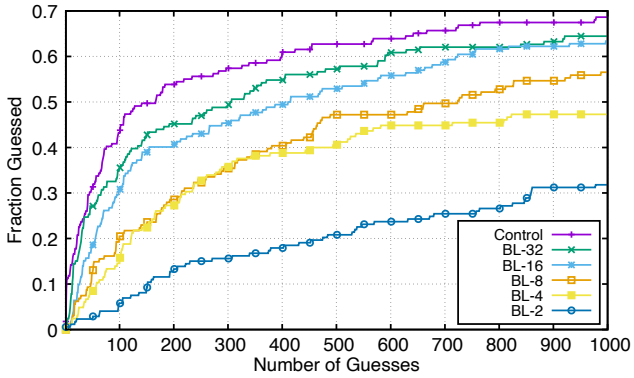
5.2 Simulated Attacker

A simulated attacker guesses a set of unknown authentications based on a set of training data. Using published data of Android patterns from von Zezschwitz et al. [30], Aviv et al. [4], Uellenbeck et al. [28], and Loge et al. [18], we first created a training data set where we ordered the patterns by their frequency of occurrence, starting from the most common patterns. Our training set consisted of a total of 4,637 patterns of which 581 are unique. In cases where multiple patterns had a similar number of occurrences, we used a Markov Model to order them based on their probability of occurrence. Using data of all possible unlock patterns from Aviv et al. [4], we trained our model to compute the transition probabilities, using Laplace smoothing to ensure no zero probability transitions existed for valid transitions not appearing in the training data. Using the Markov Model once more, we extended our initial training set by adding all other possible Android unlock patterns based on their likelihood of occurrence. Our final training data set was comprised of 389,112 patterns, the total possible number of Android unlock patterns.

Figure 4a and Figure 4b show the results of a simulated guess for up to 30 and 1000 patterns respectively. As can be seen from both graphs, blocklists reduce the fraction of patterns guessed. After 30 guesses, the simulated attacker can guess 23.7 % of patterns in the control treatment, 20.5 % in BL-32, 11.6 % in BL-16, 7.5 % in BL-8, 4.8 % in BL-4, and 2.3 % in BL-2. After 1000 guesses, the attacker can guess 68.7 % of patterns in the control group, 64.5 % in BL-32,



(a) Success rate for up to 30 guesses.



(b) Success rate for up to 1000 guesses.

Figure 4: Success rates of a simulated attacker.

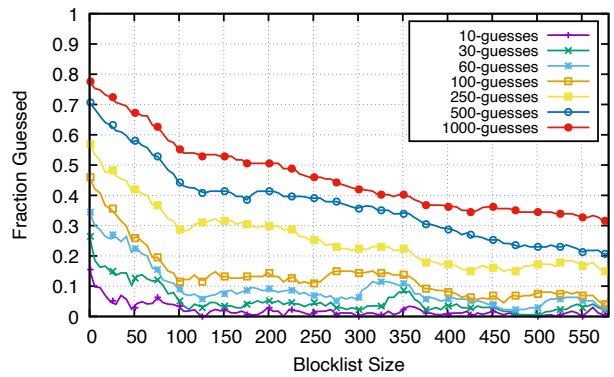


Figure 5: Effect of blocklist size on simulated guessing.

63.4 % in BL-16, 56.5 % in BL-8, 47.3 % in BL-4, and 31.8 % in BL-2. Hence, we can conclude that similar to perfect knowledge guessing, even the smallest blocklists reduce the simulated attacker's success rate.

5.3 Appropriate Blocklist Size

While a large blocklist is beneficial for security, it is important to have an appropriate blocklist size to limit negative effects on the user experience. As shown in Table 2, chances of users encountering a blocklist are higher as the blocklist size increases. Hence, we now discuss an appropriate blocklist size to balance the security and usability of unlock patterns.

Our experiment allowed us to collect not just the final patterns but also all other patterns selected by participants that were rejected due to a blocklist. With knowledge of each participant's pattern selection attempts, we simulated different blocklist sizes to determine the pattern they would have selected given a certain blocklist size. Finally, we performed a simulated guessing attack to determine the fraction of patterns that would be guessed for the different simulated blocklist sizes after a varied number of guessing attempts.

Figure 5 shows our simulation results. Initially, a lot of patterns can be guessed when there is no blocklist in place, i.e., when the blocklist size is 0. As the blocklist size increases, the fraction of patterns guessed also decreases through a series of dips and peaks, caused by participants settling again on popular patterns after encountering a blocklist warning on their previous choice. By entering the first dip for instance, the attacker is most disadvantaged as it is no longer possible to solely rely on guessing first choice patterns; but more and more second choice patterns need to be considered as well. Ultimately, the blocklist restricts all first choices and therefore, the attacker can now guess popular second choices which results in a peak.

These series of dips and peaks suggest that a properly sized blocklist should be based on one of the dips as this is where the attacker is most disadvantaged. To achieve this while having minimal effect on usability, the first dip for 30 to 60 guesses that translates to a blocklist size of about 100 patterns appears to be the most ideal. This is most similar to the BL-8 treatment which blocked 105 patterns.

6 Pattern Selection Strategies

In this section, we discuss the strategies that participants used to select patterns when encountering a blocklist. Participants were asked about both their initial strategy for selecting a pattern and how that strategy changed when they encountered a blocklist. Those who did not encounter a blocklist, were asked to imagine how they would change strategies. We qualitatively coded a random sub-sample of 309 responses (about a third of the sample space), split comparably across treatments. Two coders independently coded the responses and met to collaboratively review discrepancies until agreement was met. We settled on 11 primary codes that describe participants' selection strategies. A full description of the codes can be found in Table 5 in Appendix B.

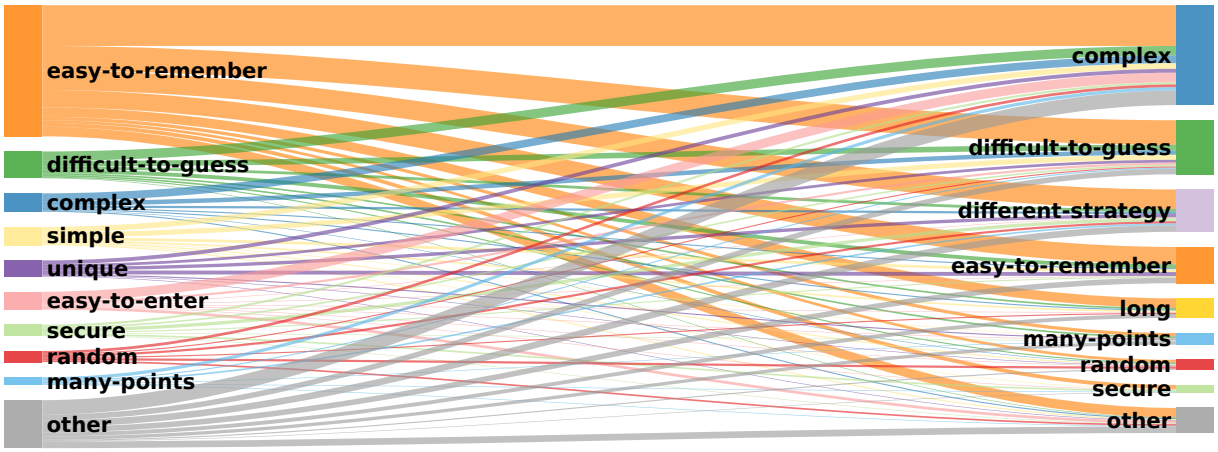


Figure 6: Changes to pattern selection strategies upon encountering a blocklist.

Initial Strategies. To understand the initial selection strategies, participants were asked about how they selected their unlock pattern, with participants that encountered a blocklist specifically asked about their strategy prior to encountering the blocklist warning. The vast majority of participants indicated selecting a pattern that would be easy to remember (70.6 %). This matches inquiries from prior work [14].

Other common strategies mentioned by participants included patterns that would be "difficult to guess" (15.1 %) and "complex" (10.0 %). There was a roughly equal split between participants who valued complexity and simplicity in their patterns, with another 10.0% of responses being tagged as "simple." 9.1 % of participants indicated that they chose their pattern to be "unique" or "uncommon" while 8.7 % mentioned selecting a pattern that would be "easy to enter".

These results indicate that prior to encountering a blocklist, most users are more concerned about selecting a pattern that would be easy to remember rather than secure. This has also been demonstrated in other studies whereby most users choose convenience over security or privacy [17].

Post-Blocklist Strategies. Participants that encountered a blocklist warning were asked how their strategies changed upon encountering the blocklist while those that did not were asked to describe how they imagine their strategies would change upon encountering such a warning. The greatest percentage of participants indicated choosing "complex" patterns (49.5 %) after encountering a blocklist, while 28.2 % changed their patterns to be "difficult to guess". A further 20.7 % of participants indicated that they would change their strategy in some way but did not specify exactly how they would do so.

Only 17.8 % of participants chose "easy-to-remember" patterns following a blocklist warning, a significant reduction compared to participants that used this strategy prior to encountering a blocklist. About 13.9 % of participants' responses were tagged as "long" or "many-points", meaning that they wanted their patterns to cover many contact points for a longer pattern, indicating a desire for complexity or a pattern that is harder to guess. A small percentage (5.5 %) of participants stated that they chose their pattern at random after the blocklist encounter.

These results show the positive security effects of blocklists, with a majority of users indicating using strategies that are more security minded, either making their patterns more complex or harder to guess.

Changes of Strategy. Figure 6 shows how participants' pattern selection strategies changed after they encountered a blocklist, with their strategies prior to a blocklist on the left and their strategies after encountering a blocklist on the right.

The most significantly changed strategies were "easy-to-remember" and "complex", with the number of participants who selected their pattern to be easy to remember decreasing by about 74.8 %. The participants who selected complex patterns increase by 393.5 %, after encountering a blocklist. The number of participants whose responses indicated security in general (complex, difficult-to-guess, long, secure) increased by 190 %. Additionally, before encountering a blocklist, only 4.2 % of participants indicated that they wanted their patterns to be long. In contrast, 13.9 % of participants increased their pattern length after encountering a blocklist warning.

Our results suggest that when users are not primed to think about the security of their patterns, they tend to prefer memorability and convenience. However, after they encountered a blocklist, the most common strategies were to make their patterns complex (49.5 %) and difficult to guess (28.2 %). This shows that blocklists can meaningfully encourage users to consider security just as much as they consider convenience when selecting patterns to secure their smartphones.

Table 4: Usage statistics for the control and the 5 blocklist treatments.

	Control	BL_32	BL_16	BL_8	BL_4	BL_2	Hit BL	No BL	Total
Mean Selection Time	13.64s	13.41s	16.67s	19.27s	25.52s	34.24s	34.50s	12.31s	20.52s
Median	7.38s	9.12s	12.34s	13.88s	17.48s	26.70s	27.99s	7.74s	13.38s
Standard Deviation	26.91s	12.17s	15.98s	17.04s	25.25s	29.23s	24.14s	18.34s	23.27s
Mean Entry Time	1.53s	1.46s	1.53s	1.73s	1.87s	1.79s	1.75s	1.59s	1.65s
Median	1.27s	1.19s	1.33s	1.46s	1.53s	1.62s	1.52s	1.32s	1.40s
Standard Deviation	1.10s	0.94s	0.83s	1.00s	1.35s	0.91s	1.04s	1.04s	1.04s
Mean Recall Attempts	1.33	1.35	1.27	1.35	1.52	1.52	1.53	1.31	1.39
Median	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Standard Deviation	0.87	0.82	0.64	0.78	1.03	1.03	1.00	0.79	0.88
Recall Success Rate	100.00%	99.55%	100.00%	99.54%	100.00%	99.62%	99.82%	99.76%	99.78%
Mean SUS Score	78.64	78.77	78.01	76.96	76.47	71.62	71.40	79.84	76.72
Median	82.5	80.0	80.0	80.0	77.5	75.0	72.5	82.5	77.5
Standard Deviation	17.37	16.51	16.47	16.84	16.80	17.82	17.74	15.95	17.12

7 Usability

In this section, we discuss the usability of patterns selected across treatments. We begin by discussing the amount of time participants took to select and enter their patterns followed by short-term recall rates across treatments. Afterwards, we discuss System Usability Score (SUS) before reporting on a series of Likert-based responses regarding usability.

Selection Time. Our study recorded the amount of time participants took to select and enter their patterns. As can be seen in Table 2, participants in the control group took on average 13.64 seconds to select a pattern, compared to 34.24 seconds on average in the largest BL-2 treatment. The 151 % increase in selection time is likely due to users encountering the blocklist multiple times as well as the extra time needed to develop more complex patterns. Among the smaller blocklist treatments, the average selection time varied by a few seconds, with participants requiring on average 5.63 more seconds to select a pattern in BL-8 compared to the control group. Using a one-way analysis of variance (ANOVA), we find significant difference ($f = 22.06$, $p < 0.05$) for selection time across treatments. By performing a post-hoc pairwise analysis (with Holm-Sidak correction), we find that the control treatment is not significantly different from the small blocklist treatments i.e. BL-32 ($p = 0.99$), BL-16 ($p = 0.99$) and BL-8 ($p = 0.76$) but significantly differs from the large blocklist groups ($p < 0.05$) i.e. BL-4 and BL-2. Further, Cohen's effect size values suggest a medium effect size between the control group and BL-4 ($d = 0.46$), but a fairly large effect size between the control group and BL-2 ($d = 0.73$). These results indicate that larger blocklists can significantly increase the time used to select a pattern, showing the need to appropriately size blocklists to preserve the usability of unlock patterns.

Entry Time. The average entry times remained mostly unaffected by the blocklists. In the control, participants took on average 1.53 seconds. The only notable changes can be seen for the large blocklist treatments where entry times rose marginally to 1.87 seconds for participants in the BL-4 treatment and 1.79 seconds for BL-2. A one-way ANOVA found significant difference ($f = 4.10$, $p < 0.05$) for entry time across treatments. However, after performing a post-hoc pairwise analysis (with Holm-Sidak correction) we do not find significant difference between any of the treatments, suggesting that blocklists have limited impact on entry time of patterns.

Recall. The vast majority of participants were able to recall their patterns later in the survey as shown in Table 4 regardless of their treatment. Recall rates, albeit short-term, were not significantly different across treatments nor between those that hit and those that did not hit a blocklist. However, the average number of attempts needed to recall their patterns did vary across treatments. In the control group, users needed 1.33 attempts while BL-2 treatment participants required 1.52 attempts on average. The users who did not hit a blocklist within any treatment took 1.31 attempts and those who did took 1.53 attempts on average. While these results suggest that patterns selected after encountering a blocklist tend to be slightly less memorable compared to those selected without a blocklist, we do not find any significant difference ($H = 9.40$, $p = 0.09$) across the treatments using the Kruskal-Wallis test. Note the recall rates captured in our experiment were short-term due to our focus on the immediate impact of blocklists; exploring long-term recall is a promising area of future work.

Usability Perceptions. We used the System Usability Scale (SUS) to measure the perceived usability of the pattern scheme in presence of different blocklists. As shown in the last row of Table 4, the SUS scores are acceptable across all

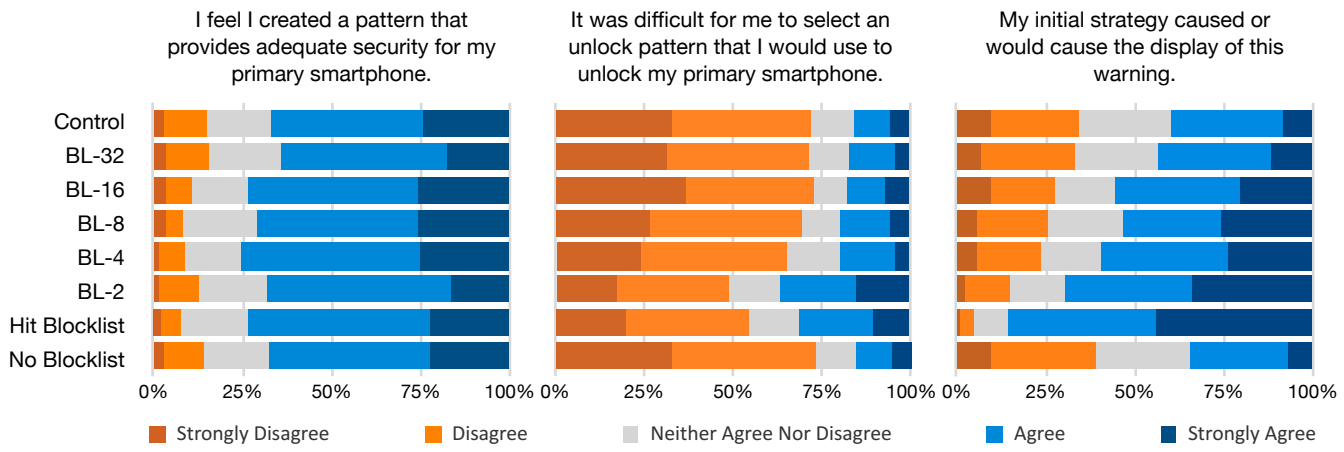


Figure 7: Agreement with Likert-scale questions relating to security and usability perceptions of unlock patterns.

the treatments. Apart from BL-2 where users recorded an SUS score of 71.6, all other blocklist treatments had SUS scores that were comparable to the control treatment, ranging from 76.5 to 78.6. We anticipate the lower but acceptable SUS score in BL-2 was due to users getting frustrated due to the large blocklist size. Therefore, it is important to use an appropriate blocklist size in order to have minimal impact on the usability of unlock patterns.

Security vs. Usability Tradeoffs. Participants were asked a series of Likert questions on the perceived security and usability of their patterns (cf. Figure 7). First, participants were asked if they felt that they "created a pattern that provides adequate security" for their primary smartphone. Across all treatments, ~70 % of participants agreed that they chose a secure pattern regardless of blocklist encounters. This could be due to social desirability bias, where participants over-report the security of their patterns to seem more favorable in a security-focused study. However, a Mann-Whitney U test showed significant difference ($U = 110674.5$, $p < 0.05$) in perceived security for participants that did and did not encounter a blocklist, suggesting that seeing a warning marginally increases ($\eta^2 = 0.003$) users' perception of the security of their patterns.

Participants were also asked if it was difficult to select an unlock pattern that they would use to unlock their primary smartphone. The percentage of participants who agreed with the statement increased as the blocklist size grew, meaning that strict blocklists made it harder for people to select usable patterns. BL-2 (20.0 %) and BL-4 (37.0 %) treatments had the largest percentage of participants agreeing with this statement. Using a Mann-Whitney U test, we find that participants that encounter a blocklist think it is more difficult ($U = 90388.5$, $p < 0.05$) to select a pattern compared to participants that do not. We anticipate that this small increase ($\eta^2 = 0.038$) is likely caused by user frustration with large blocklists, further reinforcing the need to appropriately size blocklists.

When prompted about their agreement with the statement "my initial strategy caused the display of this warning" over 80 % of participants that encountered a blocklist agreed. In contrast, less than 35 % of participants that did not hit a blocklist agreed that their initial strategy would cause the display of the warning. The control group participants were split roughly evenly with slightly more people agreeing with the statement. As the blocklist size increased, more participants agreed that their strategy caused the warning, with 70 % of BL-2 participants agreeing. A Mann-Whitney U test showed significant difference ($U = 43562.0$, $p < 0.05$) in agreement with the statement for those who encountered the blocklist versus those who did not. This suggests that after encountering a blocklist, users are 27.8 % more likely ($\eta^2 = 0.278$) to think critically about the security of their patterns

Our findings on the usability of Android patterns with blocklists seem to support our hypothesis. While large blocklists do increase users' perception of the security of their patterns, they also make patterns less memorable and less usable. Moderate blocklists such as BL-8 and BL-16 seem to improve security without the huge usability trade-off incurred by BL-2. This further shows the need to select an appropriate blocklist size to avoid user frustration during pattern selection.

8 Discussion

Android unlock patterns continue to be a popular mobile authentication mechanism, 27 % of respondents in our study use them, matching similar reported usage in prior studies [17, 19]. This makes Android patterns the second most commonly used authentication on mobile devices after PINs. However, despite their popularity, patterns are comparatively less secure than both PINs and passwords [4, 19, 28], and have remained largely unchanged since they were first launched in 2008, with no significant security updates.

Our work suggests that the usage of blocklists, even quite small in size, can have dramatic improvements on the security of user-chosen unlock patterns. The blocklist warnings also primed participants to be more security conscious of their pattern choices, and had limited impact on short-term recall and entry times. Our results indicate that a blocklist with around 100 patterns would balance the security and usability needs sufficiently and could be deployed quickly and efficiently with minimal changes to Android's existing pattern interface.

Compared to most suggestions proposed to improve the security of unlock patterns such as rearrangement of points on the grid [27], use of strength meters [2, 24, 25] or providing guidance during selection [12], blocklists require the least updates to the simple interface that makes patterns so popular. In fact, existing warnings such as the one in use on Apple iOS can easily be adapted. Further, while blocklists have already been shown to improve security on other mobile authentication schemes such as PINs [19] and Knock Codes [21], suggestions such as increasing the grid size have proven not to have meaningful security benefits [4]. While proposals such as Double Patterns [14] improve security, it remains unclear if they will be widely adopted because unlike blocklists, they alter both the selection and entry procedure of unlock patterns.

Our qualitative results demonstrated how users do not have a good sense of the security of their pattern choices, with most users (even those that selected easily guessable patterns) indicating their patterns to be secure. While this may be due to social desirability bias, we do observe that encountering a blocklist forces users to think about security of their patterns, with users resorting to patterns that are either complex or difficult to guess. In contrast, most users are primarily concerned about memorability of their patterns prior to encountering a blocklist. This suggests that the usage of blocklists can force users to consider security when selecting unlock patterns.

The biggest challenge with deployment of blocklists is asking participants to update their pattern if the one they currently use is on the blocklist. Research on password reuse notifications may be of benefit in solving this problem. For example, Golla et al. [16] investigate different notifications for password reuse which could be adapted to encourage participants to update their pattern to one not on the blocklist, including forcing a password reset. Since non-enforcing blocklists have been shown to have limited security benefits on user-chosen PINs [19], we recommend using enforcing blocklists that would force users to select patterns more diversely.

Long term memorability of patterns selected in the presence of blocklists could pose another challenge to the adoption of blocklists on Android patterns. While short-term recall times and attempts only varied marginally for participants in the blocklist treatments compared to the control group, our study design did not allow us to measure recall over an extended duration of time, which can be explored further in future work. However, similar approaches have been successfully used in prior work [4, 19, 21] in the security community. Additionally, blocklists have successfully been used on other mobile authentication schemes such as PINs [19].

While we primarily focus on a simulated and perfect knowledge attacker for our analysis, further work is needed to determine how other attackers including a targeted attacker would perform in guessing patterns, particularly if they are aware of the blocklist used. While we observe positive change of strategies from simple to complex after encountering a blocklist, this very information could further improve the guessing performance of an informed attacker. On the other hand, this change of strategy is likely to make it harder for attacks such as shoulder surfing [5].

Future work may also analyze pattern strategies used by participants in real time in order to provide more tailored blocklist warnings. For instance, if a user selects a shape such as a letter, the blocklist warning could inform the user that their pattern is a letter and can therefore be easily guessed. Other areas for more work include investigating whether different blocklists are needed for different communities. Our blocklists were constructed using common patterns observed in prior work, with these patterns primarily collected from users in Western countries. This might explain the reason for most users starting their patterns from the upper left corner as Western writing begins from the top left. Other work could also explore whether the blocklists would need to be updated over time as this was outside the scope of our study.

9 Conclusion

In this paper, we studied the security and usability of blocklists on user-selected unlock patterns, a feature currently unavailable on Android but used by Apple's iOS to improve the security of user-selected PINs. We conducted an online survey where $n = 1006$ participants selected patterns across 6 treatments: a control treatment and 5 blocklist-enforcing treatments. We find that even small blocklists improve the security of unlock patterns. For a simulated attacker that must guess patterns based on some training data, the attacker's performance is reduced from 24 % to 20 % of patterns successfully guessed after 30 guesses; the largest blocklist reduces the attacker's performance further down to only 2 % after a similar number of guessing attempts. For usability, blocklists had minimal impact on short-term recall rates and entry times, with SUS scores indicating good usability when selecting patterns even in the presence of a blocklist. From our results, we recommend a blocklist size of about 100 patterns to balance the security and usability of patterns. Adding this feature to the existing implementation of Android patterns can be done easily and does not require changes to the original interface.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1845300. This research was further supported by the research training group "Human Centered Systems Security" sponsored by the state of North Rhine-Westphalia, Germany, and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2092 CASA – 390781972.

References

[1] Daniel Amitay. Most Common iPhone Passcodes, June 2011. http://danielamitay.com/blog/2011/6/13/most-common-iphone-passcodes, as of July 6, 2021.

[2] Panagiotis Andriotis, Theo Tryfonas, and George Oikonomou. Complexity Metrics and User Strength Perceptions of the Pattern-Lock Graphical Authentication Method. In *Conference on Human Aspects of Information Security, Privacy and Trust*, HAS '14, pages 115–126, Heraklion, Crete, Greece, June 2014. Springer.

[3] Panagiotis Andriotis, Theo Tryfonas, George Oikonomou, and Can Yildiz. A Pilot Study on the Security of Pattern Screen-Lock Methods and Soft Side Channel Attacks. In *ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '13, pages 1–6, Budapest, Hungary, April 2013. ACM.

[4] Adam J. Aviv, Devon Budzitowski, and Ravi Kuber. Is Bigger Better? Comparing User-Generated Passwords on 3x3 vs. 4x4 Grid Sizes for Android's Pattern Unlock. In *Annual Computer Security Applications Conference*, ACSAC '15, pages 301–310, Los Angeles, California, USA, December 2015. ACM.

[5] Adam J. Aviv, John T. Davin, Flynn Wolf, and Ravi Kuber. Towards Baselines for Shoulder Surfing on Mobile Authentication. In *Annual Conference on Computer Security Applications*, ACSAC '17, pages 486–498, Orlando, Florida, USA, December 2017. ACM.

[6] Adam J. Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M. Smith. Smudge Attacks on Smartphone Touch Screens. In *USENIX Workshop on Offensive Technologies*, WOOT '10, pages 1–7, Washington, District of Columbia, USA, August 2010. USENIX.

[7] Adam J. Aviv, Benjamin Sapp, Matt Blaze, and Jonathan M. Smith. Practicality of Accelerometer Side Channels on Smartphones. In *Annual Computer Security Applications Conference*, ACSAC '12, pages 41–50, Orlando, Florida, USA, December 2012. ACM.

[8] Adam J. Aviv, Flynn Wolf, and Ravi Kuber. Comparing Video Based Shoulder Surfing with Live Simulation and Towards Baselines for Shoulder Surfing on Mobile Authentication. In *Annual Conference on Computer Security Applications*, ACSAC '18, pages 453–466, San Juan, Puerto Rico, USA, December 2018. ACM.

[9] Joseph Bonneau. The Science of Guessing: Analyzing an Anonymized Corpus of 70 Million Passwords. In *IEEE Symposium on Security and Privacy*, SP '12, pages 538–552, San Jose, California, USA, May 2012. IEEE.

[10] Joseph Bonneau, Sören Preibusch, and Ross Anderson. A Birthday Present Every Eleven Wallets? The Security of Customer-Chosen Banking PINs. In *Financial Cryptography and Data Security*, FC '12, pages 25–40, Kralendijk, Bonaire, February 2012. Springer.

[11] Seunghun Cha, Sungsu Kwag, Hyoungshick Kim, and Jun Ho Huh. Boosting the Guessing Attack Performance on Android Lock Patterns with Smudge Attacks. In *ACM Asia Conference on Computer and Communications Security*, ASIA CCS '17, pages 313–326, Abu Dhabi, United Arab Emirates, April 2017. ACM.

[12] Geumhwan Cho, Jun Ho Huh, Junsung Cho, Seongyeol Oh, Youngbae Song, and Hyoungshick Kim. SysPal: System-Guided Pattern Locks for Android. In *IEEE Symposium on Security and Privacy*, SP '17, pages 338–356, San Jose, California, USA, May 2017. IEEE.

[13] Alexander De Luca, Marian Harbach, Emanuel von Zezschwitz, Max-Emanuel Maurer, Bernhard Ewald Slawik, Heinrich Hussmann, and Matthew Smith. Now You See Me, Now You Don't: Protecting Smartphone Authentication from Shoulder Surfers. In *ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2937–2946, Toronto, Ontario, Canada, April 2014. ACM.

[14] Tim Forman and Adam J. Aviv. Double Patterns: A Usable Solution to Increase the Security of Android Unlock Patterns. In *Annual Conference on Computer Security Applications*, ACSAC '20, pages 219–233, Virtual Conference, December 2020. ACM.

[15] Maximilian Golla, Jan Rimkus, Adam J. Aviv, and Markus Dürmuth. Work in Progress: On the In-Accuracy and Influence of Android Pattern Strength Meters. In *Workshop on Usable Security and Privacy*, USEC '19, San Diego, California, USA, February 2019. ISOC.

[16] Maximilian Golla, Miranda Wei, Juliette Hainline, Lydia Filipe, Markus Dürmuth, Elissa Redmiles, and Blase Ur. "what was that site doing with my facebook password?": Designing password-reuse notifications. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, CCS '18, page 1549–1566, New York, NY, USA, 2018. Association for Computing Machinery.

[17] Marian Harbach, Emanuel von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 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 '14, pages 213–230, Menlo Park, California, USA, July 2014. USENIX.

[18] Marte Løge, Markus Dürmuth, and Lillian Røstad. On User Choice for Android Unlock Patterns. In *European Workshop on Usable Security*, EuroUSEC '16, Darmstadt, Germany, July 2016. ISOC.

[19] Philipp Markert, Daniel V. Bailey, Maximilian Golla, Markus Dürmuth, and Adam J. Aviv. This PIN Can Be Easily Guessed: Analyzing the Security of Smartphone Unlock PINs. In *IEEE Symposium on Security and Privacy*, SP '20, pages 286–303, San Francisco, California, USA, May 2020. IEEE.

[20] Elissa M. Redmiles, Yasemin Acar, Sascha Fahl, and Michelle L. Mazurek. A Summary of Survey Methodology Best Practices for Security and Privacy Researchers. Technical Report CS-TR-5055, UM Computer Science Department, May 2017.

[21] Raina Samuel, Philipp Markert, Adam J. Aviv, and Iulian Neamtiu. Knock, Knock. Who's There? On the Security of LG's Knock Codes. In *Symposium on Usable Privacy and Security*, SOUPS '20, pages 37–59, Virtual Conference, August 2020. ACM.

[22] Florian Schaub, Ruben Deyhle, and Michael Weber. Password Entry Usability and Shoulder Surfing Susceptibility on Different Smartphone Platforms. In *International Conference on Mobile and Ubiquitous Multimedia*, MUM '12, pages 13:1–13:10, Ulm, Germany, December 2012. ACM.

[23] Stefan Schneegass, Frank Steimle, Andreas Bulling, Florian Alt, and Albrecht Schmidt. Smudgesafe: Geometric Image Transformations for Smudge-Resistant User Authentication. In *Conference on Ubiquitous Computing*, UbiComp '14, pages 775–786, Seattle, Washington, USA, September 2014. ACM.

[24] Youngbae Song, Geumhwan Cho, Seongyeol Oh, Hyoungshick Kim, and Jun Ho Huh. On the Effectiveness of Pattern Lock Strength Meters: Measuring the Strength of Real World Pattern Locks. In *ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2343–2352, Seoul, Republic of Korea, April 2015. ACM.

[25] Chen Sun, Yang Wang, and Jun Zheng. Dissecting Pattern Unlock: The Effect of Pattern Strength Meter on Pattern Selection. *Journal of Information Security and Applications*, 19(4–5):308–320, November 2014.

[26] Hai Tao and Carlisle Adams. Pass-Go: A Proposal to Improve the Usability of Graphical Passwords. *International Journal of Network Security*, 7(2):273–292, September 2008.

[27] Harshal Tupsamudre, Vijayanand Banahatti, Sachin Lodha, and Ketan Vyas. Pass-O: A Proposal to Improve the Security of Pattern Unlock Scheme. In *ACM Asia Conference on Computer and Communications Security*, ASIA CCS '17, pages 400–407, Abu Dhabi, United Arab Emirates, April 2017. ACM.

[28] Sebastian Uellenbeck, Markus Dürmuth, Christopher Wolf, and Thorsten Holz. Quantifying the Security of Graphical Passwords: The Case of Android Unlock Patterns. In *ACM Conference on Computer and Communications Security*, CCS '13, pages 161–172, Berlin, Germany, October 2013. ACM.

[29] Emanuel von Zezschwitz, Alexander De Luca, Philipp Janssen, and Heinrich Hussmann. Easy to Draw, but Hard to Trace?: On the Observability of Grid-based (Un)Lock Patterns. In *ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2339–2342, Seoul, Republic of Korea, April 2015. ACM.

[30] Emanuel von Zezschwitz, Malin Eiband, Daniel Buschek, Sascha Oberhuber, Alexander De Luca, Florian Alt, and Heinrich Hussmann. On Quantifying the Effective Passsword Space of Grid-Based Unlock Gestures. In *Conference on Mobile and Ubiquitous Multimedia*, MUM '16, pages 201–212, Rovaniemi, Finland, December 2016. ACM.

[31] Ding Wang, Qianchen Gu, Xinyi Huang, and Ping Wang. Understanding Human-Chosen PINs: Characteristics, Distribution and Security. In *ACM Asia Conference on Computer and Communications Security*, ASIA CCS '17, pages 372–385, Abu Dhabi, United Arab Emirates, April 2017. ACM.

[32] Guixin Ye, Zhanyong Tang, Dingyi Fang, Xiaojiang Chen, Willy Wolff, Adam J. Aviv, and Zheng Wang. A Video-based Attack for Android Pattern Lock. *ACM Transactions on Privacy and Security*, 21(4):19:1–19:31, July 2018.

Appendix

A Survey Material

Purpose of Study and Task Description
You are being asked to participate in a research study focused on the effectiveness of mobile authentication on an Android device. Androids implement pattern locks rather than traditional security parameters, for example, numeric PINs or alphanumeric passwords.

You will be asked to complete a short survey that requires you to generate a set of Android patterns under a security scenario, such as locking your device. Your eventual choices will be used in the final evaluation, as well as your responses to a set of security and usability questions.

The expected completion time of the survey is 8–10 minutes, and no more than 1 hour. You will be compensated $1.00 for your participation.

Device Usage Questions
When referring to "mobile devices" throughout this survey, consider these to include smartphones and tablet computers. Traditional laptop computers, two-in-one computers, like the Microsoft Surface, or e-readers, like the Amazon Kindle, are not considered mobile devices for the purposes of this survey.

1. How many mobile devices do you use regularly?
   ○ 0 ○ 1 ○ 2 ○ 3 ○ 4+

2. What brands of smartphone do you use for personal use? (Select all that apply)
   ☐ Apple ☐ Samsung ☐ LG ☐ Motorola ☐ Google/Pixel/Nexus ☐ Huawei ☐ ZTE ☐ Other

3. What biometric method do you use most often to unlock your primary personal smartphone?
   ○ I do not use a biometric ○ Fingerprint ○ Face ○ Iris ○ Other Biometric ○ I do not use a smartphone ○ Prefer Not to Say

If participants indicated to use a biometric:

4a. You have indicated that you use a biometric on your smartphone. Please answer the following question related to your response. How do you unlock your primary personal smartphone when you reboot the device or if your biometric fails?
○ Pattern Unlock ○ 4-Digit PIN ○ 6-Digit PIN ○ PIN of other length ○ Alphanumeric Password ○ I use an unlock method not listed ○ I do not use a smartphone ○ Prefer Not to Say

If participants indicated not to use a biometric:

4b. You have indicated that you do not use a biometric on your smartphone. Please answer the following question related to your response. What unlock method do you use on your primary personal smartphone?
○ Pattern Unlock ○ 4-Digit PIN ○ 6-Digit PIN ○ PIN of other length ○ Alphanumeric Password ○ I use an unlock method not listed ○ I do not use a smartphone ○ Prefer Not to Say

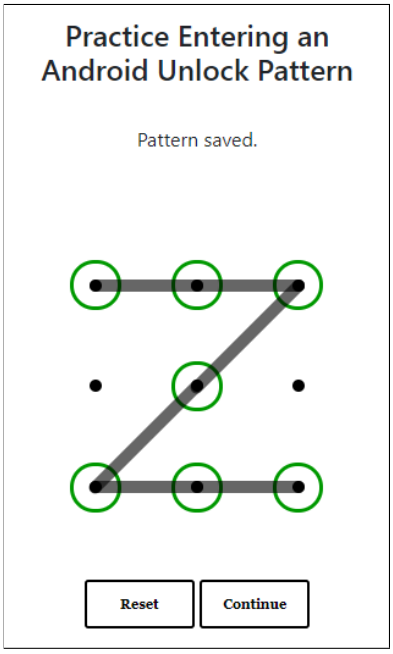


Figure 8: User interface for entering patterns.

What are Android Pattern Locks?
Pattern Locks are used to unlock your smartphone, like a PIN. Patterns require you to “draw” a shape that connects at least four of the contact points without lifting your finger or repeating a contact point. Displayed below is the Pattern Lock interface on a Samsung Android mobile device.

A Little Bit of Practice
On the next page, you will have a chance to practice entering an Android unlock pattern before proceeding with the rest of this survey, where we will ask you to select your own pattern that you would utilize on your primary smartphone.

Practice Entering an Android Unlock Pattern

Interface as shown in Figure 8

Instructions
For this survey, you will be asked to create an Android unlock pattern you would likely use to secure your primary smartphone. You will need to recall this unlock pattern later in the survey, so choose something that is secure and memorable as you may use on your primary smartphone. We ask that you DO NOT write down your patterns or use other aids to help you remember.
I understand that I should not write down my unlock pattern or use other aids to assist in the survey. ○ I understand

I understand that I will be asked to create an unlock pattern that I would use on my primary smartphone. ○ I understand

Selection

Interface as shown in Figure 8

Simple Usability Scale

Select your agreement/disagreement with the following statements. Please note that the term "system" refers to the selection of the Android unlock pattern.

5. I think that I would like to use this system frequently.
   ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
6. I found the system unnecessarily complex.
   ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
7. I thought the system was easy to use.
   ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
8. I think that I would need the support of a technical person to be able person to be able to use this system.
   ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
9. I thought there was too much inconsistency in this system.
   ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
10. I found the various functions in this system were well integrated.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
11. I would imagine that most people would learn to use this system very quickly.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
12. Select Agree as the answer to this question.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
13. I found this system very cumbersome to use.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
14. I felt very confident using this system.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
15. I needed to learn a lot of things before I could get going with this system.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

Thinking about the Android unlock pattern you just chose:

16. I feel I created an Android unlock pattern that provides adequate security for unlocking my primary smartphone.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree
17. It was difficult for me to select an Android unlock pattern that I would use to unlock my primary smartphone.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

For participants who received a blocklist warning:

We noticed that you received the following warning while choosing your pattern:

Warning as shown in Figure 1

18a. Prior to seeing the warning above, what was your strategy for choosing your unlock pattern? [Open Text]

19a. After receiving the warning message, please describe how or if your strategy changed when choosing your unlock pattern. [Open Text]

20a. My initial strategy caused the display of this warning.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

For participants who did not receive a blocklist warning:

18b. What was your strategy when choosing your Android unlock pattern? [Open Text]

Imagine you received the following warning message after choosing your pattern:

Warning as shown in Figure 1

19b. Please describe how or if your strategy would change as a result of the message. [Open Text]

20b. My strategy would cause this warning message to appear.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

Recall Android Pattern

Recall the Android Unlock Pattern you created previously to secure your Primary Smartphone.

Interface as shown in Figure 8

Security Comparison

Select your agreement/disagreement with the following statements.

Questions 21-24 were shown in randomized order.

21. Unlock patterns are a secure way to unlock my primary smartphone.
    ○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

22. Unlock patterns are more secure than alphanumeric passwords for unlocking my primary smartphone.
○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

23. Unlock patterns are more secure than 4-digit PIN codes for unlocking my primary smartphone.
○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

24. Unlock patterns are more secure than 6-digit PIN codes for unlocking my primary smartphone.
○ Strongly Agree ○ Agree ○ Neither Agree Nor Disagree ○ Disagree ○ Strongly Disagree

Use Unlock Pattern from Survey

25. If you were using an unlock pattern on your primary smartphone, would you use the unlock pattern you selected in this survey, or would you select a different one?
○ Yes, I would use the unlock pattern I created here on my primary smartphone.
○ No, I would not use the unlock pattern I created here and instead create a new one to use on my personal device.
○ Unsure, I may or may not use the unlock pattern I created here on my personal device.

26. [You have indicated that you would use / You have indicated that you are unsure if you / You have indicated that you would not use if you would use] the unlock pattern that you created in this survey on your personal mobile device. Please expand on why you [would / are unsure if you would / would not] use the unlock pattern you created here. [Open Text]

Demographics

Please enter your demographic information.

27. Select your age:
○ 18-24 ○ 25-29 ○ 30-34 ○ 35-39 ○ 40-44 ○ 45-49 ○ 50-54 ○ 55-59 ○ 60-64 ○ 65+ ○ Prefer Not to Say

28. With which gender do you most identify?
○ Female ○ Male ○ Non-Binary/Third Gender ○ Not Described Here ○ Prefer Not to Say

29. What is your dominant hand?
○ Left Handed ○ Right Handed ○ Ambidextrous ○ Prefer Not to Say

30. Where you live is best described as
○ Urban ○ Suburban ○ Rural ○ Prefer Not to Say

31. What is the shape of a red ball?
○ Red ○ Blue ○ Square ○ Round ○ Prefer Not to Say

32. What is the highest degree or level of school you have completed?
○ Some high school ○ High school ○ Some college ○ Trade, technical, or vocational training ○ Associate's Degree ○ Bachelor's Degree ○ Master's Degree ○ Professional degree ○ Doctorate ○ Prefer Not to Say

33. Which of the following best describes your educational background or job field?
○ I have an education in, or work in, the field of computer science, computer engineering or IT.
○ I do not have an education in, nor do I work in, the field of computer science, computer engineering or IT.
○ Prefer Not to Say

One More Thing...

Please indicate if you've honestly participated in this survey and followed instructions completely. You will not be penalized/rejected for indicating 'No' but your data may not be included in the analysis:
○ Yes ○ No

B Additional Figures and Tables

Table 5: Codebook *Pattern Select Strategy*: "Prior to seeing the warning above, what was your strategy for choosing your unlock pattern?"

Code	Frequency	Sample Quote
easy-to-remember	218	"I wanted to pick a pattern that I knew I would be able to remember."
difficult-to-guess	47	"I just started drawing something that I didn't think someone would be able to guess."
complex	31	"I tried to make a somewhat complicated pattern that I could remember."
simple	31	"Something extremely basic that I've not personally used prior to this."
unique	28	"Try to get a pattern that wasn't used a lot."
easy-to-enter	27	"I chose a pattern that would be quick and easy to use everyday."
secure	19	"I wanted something that would feel secure to lock my phone."
random	17	"I just made a random pattern that came to my mind."
many-points	13	"I tried to think of a pattern that used as many dots as possible."

* Note that each quote can be assigned multiple codes.

Table 6: Codebook *Post-Blocklist Strategy* "After receiving the warning message, please describe how or if your strategy changed when choosing your unlock pattern."

Code	Frequency	Sample Quote
complex	153	"Yes I changed it to include diagonals in a more complex manner."
difficult-to-guess	87	"Choosing a pattern that I think others would be less likely to use or guess."
different-strategy	64	"I would choose a different pattern."
easy-to-remember	55	"Didn't want to create something I'd forget quick."
long	28	"I would choose a longer one."
random	16	"I tried to think of an extremely random pattern. Something that a lot of people wouldn't select."
many-points	15	"I would use more points of the grid."
secure	13	"Make it secure."

* Note that each quote can be assigned multiple codes.

Table 7: Usage of devices and unlock methods.

	Male		Female		Other		Total	
	No.	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Number of Devices	624	62 %	367	36 %	15	1 %	1006	100 %
0	0	0 %	1	0 %	0	0 %	1	0 %
1	369	37 %	225	22 %	9	1 %	603	60 %
2	213	21 %	112	11 %	6	1 %	331	33 %
3	33	3 %	24	2 %	0	0 %	57	6 %
4+	9	1 %	5	0 %	0	0 %	14	1 %
Device Brand	624	62 %	367	36 %	15	1 %	1006	100 %
Apple	86	9 %	72	7 %	2	0 %	160	16 %
Samsung	231	23 %	147	15 %	6	1 %	384	38 %
LG	37	4 %	26	3 %	1	0 %	64	6 %
Motorola	39	4 %	29	3 %	0	0 %	68	7 %
Google	57	6 %	20	2 %	1	0 %	78	8 %
Huawei	9	1 %	2	0 %	0	0 %	11	1 %
ZTE	4	0 %	1	0 %	0	0 %	5	0 %
Other	161	16 %	70	7 %	5	0 %	236	23 %
None	0	0 %	0	0 %	0	0 %	0	0 %
Biometric Method	624	62 %	367	36 %	15	1 %	1006	100 %
No biometric	152	15 %	117	12 %	4	0 %	273	27 %
Fingerprint	387	38 %	183	18 %	5	0 %	575	57 %
Face	72	7 %	49	5 %	2	0 %	123	12 %
Iris	3	0 %	1	0 %	0	0 %	4	0 %
Other	7	1 %	12	1 %	0	0 %	19	2 %
No smartphone	0	0 %	0	0 %	0	0 %	0	0 %
Prefer not to say	3	0 %	5	0 %	4	0 %	12	1 %
Unlock Method	624	62 %	367	36 %	15	1 %	1006	100 %
Pattern unlock	165	16 %	95	9 %	8	1 %	268	27 %
4-Digit PIN	289	29 %	166	17 %	3	0 %	458	46 %
6-Digit PIN	100	10 %	62	6 %	3	0 %	165	16 %
Other PIN	9	1 %	8	1 %	0	0 %	17	2 %
Alphanumeric Password	33	3 %	6	1 %	0	0 %	39	4 %
Other	23	2 %	23	2 %	0	0 %	46	5 %
No smartphone	0	0 %	0	0 %	0	0 %	0	0 %
Prefer not to say	5	0 %	7	1 %	1	0 %	13	1 %