I Don't Even Have to Bother Them!

Using Social Media to Automate the Authentication Ceremony in Secure Messaging

Elham Vaziripour, Devon Howard, Jake Tyler, Mark O'Neill, Justin Wu, Kent Seamons, Daniel Zappala
Brigham Young University
[elhamvaziripour,devonhoward,jrtyler,mto,justinwu]@byu.edu,[seamons,zappala]@cs.byu.edu

ABSTRACT

The privacy guaranteed by secure messaging applications relies on users completing an authentication ceremony to verify they are using the proper encryption keys. We examine the feasibility of social authentication, which partially automates the ceremony using social media accounts. We implemented social authentication in Signal and conducted a within-subject user study with 42 participants to compare this with existing methods. To generalize our results, we conducted a Mechanical Turk survey involving 421 respondents. Our results show that users found social authentication to be convenient and fast. They particularly liked verifying keys asynchronously, and viewing social media profiles naturally coincided with how participants thought of verification. However, some participants reacted negatively to integrating social media with Signal, primarily because they distrust social media services. Overall, automating the authentication ceremony and distributing trust with additional service providers is promising, but this infrastructure needs to be more trusted than social media companies.

CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**.

KEYWORDS

secure messaging applications, authentication ceremony, social authentication

ACM Reference Format:
Elham Vaziripour, Devon Howard, Jake Tyler, Mark O'Neill, Justin Wu, Kent Seamons, Daniel Zappala . 2019. I Don't Even Have to Bother Them!: Using Social Media to Automate the Authentication Ceremony in Secure Messaging. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300323

CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300323

1 INTRODUCTION

Secure messaging applications provide end-to-end encrypted conversations, which is particularly important where people are vulnerable to government surveillance and lack free speech protection. However, the security of these applications is dependent on users verifying the encryption keys used by the application. Most secure messaging applications provide a user interface for users to do this, using a method called the *authentication ceremony*. This typically involves users needing to scan a QR code, if they are in the same location, or making a phone call and comparing the key fingerprint verbally. Unfortunately, research has shown that users are not able to find or perform the authentication ceremony in current secure messaging applications [1, 4, 7, 12, 13].

Two recent studies have sought to improve the usability of the authentication ceremony. First, Vaziripour et al. showed that users could complete the authentication ceremony in several applications—WhatsApp, Facebook Messenger, and Viber—but only if they had instruction about potential threats and about the importance of the authentication ceremony [13]. Second, in a follow-up paper, Vaziripour et al. redesigned the authentication ceremony in Signal using opinionated design and nudges, and they showed that with this new design most users were able to find and use the authentication ceremony [14]. However, this work paid users a small monetary incentive to encourage them to have a security mindset, and it is not clear if this would carry over to situations outside of a lab study. In addition, users still had to compare a long key fingerprint over a phone call, something they considered annoying.

In this paper, we examine whether it is possible to simplify and partially automate the authentication ceremony using a method we call *social authentication*. In this method, we use social media to distribute the public keys used by Signal and automatically compare them once users have verified that we have linked the correct social media accounts for them. Our goal is to examine whether social authentication is a

feasible authentication method – whether users can perform it quickly, like using it, trust it, and understand it. We also want to determine how this method compares against the current methods of scanning a QR code or using a phone call in terms of usability, trust, and user preference. To the best of our knowledge, no prior work has explored automating the authentication ceremony.

To evaluate the effectiveness and the feasibility of social authentication, we integrated it into the Signal messaging application and conducted two studies. We first conducted a lab user study with 21 pairs of participants (42 total), with each pair trying all three methods of authentication and comparing them. We then conducted a survey of 421 participants on Mechanical Turk, to study the feasibility of social authentication and whether our lab user study results generalize to a larger population. We believe our results generalize to other secure messaging applications that also use an authentication ceremony, such as WhatsApp, Facebook Messenger, Telegram, and Viber.

Overall, we find that social authentication is considered by participants to be usable, fast, and convenient. However, participants did not understand how it helped protect the privacy of their conversations, worried about account compromise, and did not trust that social media could be used for this purpose. It is difficult for people to conceive of using a public social media account to somehow improve their privacy, and social media companies do not have a strong reputation for trustworthiness due to frequent hacking. Participants also identified trade-offs with the QR code and phone call methods, leaving no clear winner. Using results from participant preferences, we identify principles for an ideal authentication ceremony and make recommendations for future work. Our work demonstrates that there is promise for automating the authentication ceremony, but that additional work is needed to realize this goal.

2 BACKGROUND AND RELATED WORK

The effective security provided by secure messaging applications depends heavily on users completing an authentication ceremony, the process of manually verifying the fingerprints of the encryption keys being used. The authentication ceremony is essential because a hacker or the service provider could replace the keys with substitutes they choose, enabling them to decrypt a conversation's traffic if they can also intercept it. To make sure their conversations are private, users need to compare the fingerprint of their keys to ensure they match. This ensures that they have each received the correct public key for the other party.

Each application provides a different user interface to facilitate the authentication ceremony. Users may verify the public key of their partner by comparing a key fingerprint that has been encoded into Short Authentication String (SAS), a hexadecimal version of the fingerprint, or by scanning a QR code that encodes the fingerprint. Recent research shows that word-based and sentence-based encodings of a key fingerprint are more resistant to attacks and rated high on usability as compared to hexadecimal, alphanumeric, or pure numeric representations [3]. Other work shows that graphical representations are more susceptible to attack but are easy and quick to use [11]. Related work has also been done with crypto phones, which require users to compare a short checksum and verify the other user's voice. Shirvanian et al. show that users are susceptible to impersonation attacks in the checksum verification process. [8, 10]. Notably, they found that security decreases when moving from a 2-word to a 4-word checksum.

The evidence to date suggests users are unable to perform the authentication ceremony successfully in current messaging applications [1, 4, 7]. Most of the related work focuses on showing that current designs are not usable enough. Assal et al. found that only 30% of participants successfully completed the QR code verification with no errors, mainly due to the ambiguity of the connection between the key and QR code [1]. Schröder et al. show a high failure rate due to usability problems and incomplete mental models [7]. Herzberg and Leibowitz proposed that with some instruction about the ceremony itself, users can successfully find and conduct the authentication [4]. Shirvanian et al. studied key verification performance by users performing authentication on remote and local conversation partners, showing that users perform poorly under most key verification methods, especially in the remote case [9]. Vaziripour et al. showed that the success rate increases to as high as 90% by providing instruction on the necessity of the authentication ceremony in addition to potential threats [13].

One recent paper focuses on improving the design of authentication ceremony within secure messaging applications. Vaziripour et al. employed usability principles and opinionated design to develop a redesigned authentication ceremony in Signal [14]. Their modifications led to a 90% success rate in completing the authentication ceremony, as compared to 30% for the original design, and fast authentication times. This was done with a small monetary incentive that encouraged participants to be security-minded, but no direct instruction on the ceremony. However, results showed that users still did not understand the meaning or purpose of the ceremony.

Other work has explored the use of social authentication for secure email, using the Keybase service. This service has users create an account at Keybase.io, then prove ownership of their social media and other accounts (e.g. GitHub). Another user can then access Keybase, download proofs of ownership, and link a public key to the user's identity. Atwater et al. used a simulated version of Keybase to implement effective key management for secure emails [2]. Lerner et al.

also implemented a more usable secure email by integrating Keybase with the Mailvelope browser plugin [5]. To compromise a user of Keybase, an attacker would need to take control of all the accounts a user has verified with Keybase and create new proofs.

3 SOCIAL AUTHENTICATION DESIGN

In Signal, safety numbers are derived from the public keys of the communicating parties, their phone numbers, and the first message they exchange. In current versions of Signal, the authentication ceremony requires users to either scan a QR code that encodes the safety number (meaning they must be in the same location), or make a phone call outside of the Signal application to compare their safety numbers and ensure they match.

In this section we describe social authentication and how we integrated it into Signal in order to compare it with these existing methods used for the authentication ceremony.

Social Authentication

The idea behind social authentication is to post the public keys of users to their social media accounts. This provides a secondary exchange method, not controlled by Signal, for verifying that the keys (and thus the safety numbers) match. This also provides a possibility for automating the comparison of safety numbers and instead places users in a more familiar situation of identifying their contacts through social media accounts. Since users generally don't know about public key cryptography [15], this method could provide an authentication process that is more intuitive and rooted in their experiences with social media.

If we could reliably infer a user's social media accounts from their phone number, then this process could be entirely automated, but we must instead (a) have each person tell Signal which social media accounts they own, and (b) have users verify that the social media accounts shown by Signal are indeed the accounts for the person they are contacting.

Thus, our social authentication method comprises three steps. First, users register their social media accounts with the Signal app when they first install the application or first turn on the social authentication feature. The Signal app posts their public key to their social media accounts at this time. Second, the app will periodically collect the public keys of all their friends. Third, when two users start a conversation, the Signal app compares the public key of the other party to all of the public keys downloaded from social media accounts. If there is a match, it is presented to the user for confirmation that this is the right person.

The social authentication system compares the public key advertised by the Signal server to the public keys advertised individually by all of the social media networks. Thus the only way to attack the system is to compromise the Signal server and *all* of the social media networks, or to have all of these parties collude. Impersonation of one or multiple social accounts will cause a warning and inform the users that the keys do not match, but the attack is foiled.

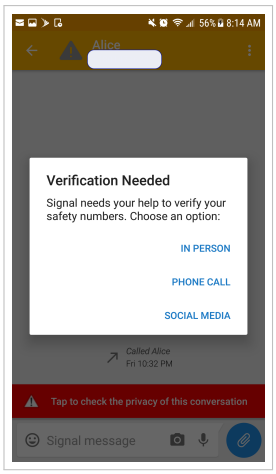
We incorporated this method of social authentication into Signal to demonstrate that it is feasible. However, creating an application that posts on multiple social media accounts requires individual approval from social media companies, and they typically review and approve professional applications, rather than research prototypes. It proved difficult and time-consuming to convince many companies to approve our app. As a result, we made several simplifications in our application that make a user study possible. We don't publish any public keys on social media accounts. This avoids the need for additional review and approval, and also avoids any potential harm that could come to users during the study. Instead, we stored the public keys in a separate server that maps each user's social media account identifier to the public key they use, as if it had been posted, and their phone number. The Signal app then queries the server for this data in order to show the user matching social media profiles. We made sure that the user interface did not change from the idealized version of social authentication to the centralized version we implemented.

Integration with Signal

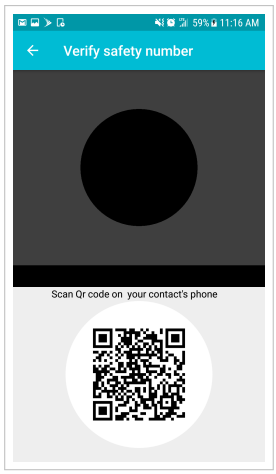
We integrated social authentication into Signal by starting with the redesigned prompts and authentication ceremony from prior work from Vaziripour et al. [14]. As shown in Figure 1a, users are encouraged to start the authentication ceremony by a red bar shown at the bottom of their conversation. Once they tap this bar, the message and options shown appear. We modified this dialog to add an option for SOCIAL MEDIA. When a user successfully completes any version of the authentication ceremony, a check mark appears next to the contact name and the red bar at the bottom of the conversation changes to blue, with a message indicating the contact has been verified. We performed a cognitive walk-through on the modified application to make sure the language used in the interface was clear.

In Person Authentication. If a user clicks the IN PERSON option, they are taken to the screen shown in Figure 1b, with the camera, activated. This screen allows a user to scan the other person's QR code, and also shows their own QR code in case the other person is the one doing the scanning. If verification is completed successfully, the app shows a green check mark and a message indicating that conversation is private. This ceremony is unchanged from Vaziripour et al.

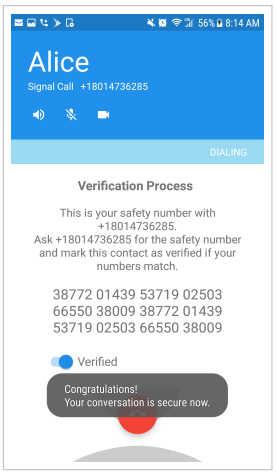
Phone Call Authentication. If a user clicks on the PHONE CALL option, they are shown a message indicating that they



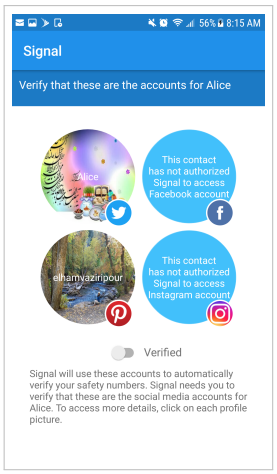
(a) Authentication ceremony options



(b) IN PERSON option

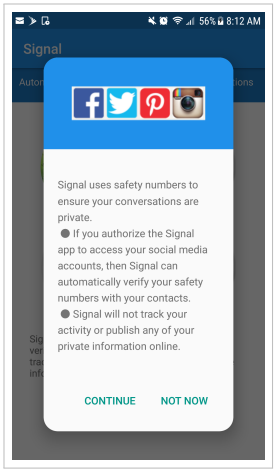


(c) PHONE CALL option

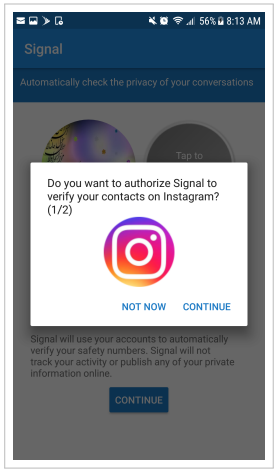


(d) SOCIAL MEDIA option

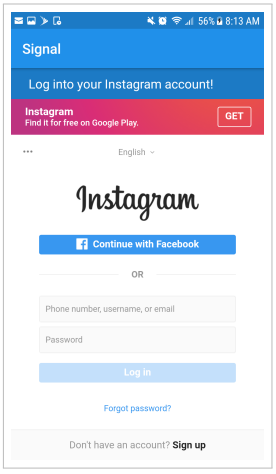
Figure 1: User interface for the authentication ceremonies



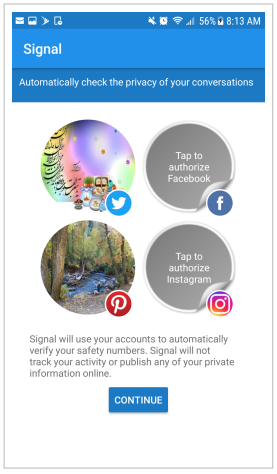
(a) Start of social authentication registration



(b) Authorization to register a social media account



(c) Login with a social media account



(d) Social media accounts verified

Figure 2: Registration phase for social authentication

will make a free phone call using the Signal app. After initiating the phone call, users see their safety number with a very brief instruction, shown in Figure 1c. Users are expected to read their safety numbers and ensure they have an identical sequence of numbers. Users are expected to switch the *Verified* toggle after they confirm that the safety numbers match. After they do this, a message appears, showing that the authentication was successful. This ceremony is also unchanged from Vaziripour et al.

Social Authentication. To use social media authentication, users first complete a registration phase when the application is initially installed. In the registration phase, users are prompted to authorize the application to access their social media accounts to verify the safety numbers automatically. If users choose to proceed with the instruction which is shown in Figure 2, they are taken to a screen showing four different social networks – Twitter, Facebook, Pinterest, and Instagram. If they click on an icon for a network, they are taken

to a screen where they can log in to that account (Figure 2b, 2c). Users can skip any of the social networks, although the application recommends they authorize as many as they can. For each of the social media accounts users authorize they will see their profile pictures as shown in Figure 2d.

Later, if a user clicks on the SOCIAL MEDIA option in Figure 1a, they are shown the screen in Figure 1d. Users are supposed to confirm that the social media profiles shown match the user they are contacting. If they match, they should switch the *Verified* toggle. If the user indicates the profiles match, then the app verifies the safety numbers posted to the social networks match the safety number that the Signal server delivered for the user. If the numbers match, the app displays a message indicating that the safety numbers match.

4 METHODOLOGY

We first conducted a lab user study to compare social authentication to in person and phone call authentication methods, then conducted a Mechanical Turk survey to examine the feasibility of social authentication and to generalize the user study results.

Lab User Study

We conducted an IRB-approved, within-subject user study, comparing different authentication ceremonies in Signal. These three methods of authentication include (1) scanning a QR code, (2) comparing a cryptographic fingerprint over a phone call, and (3) our new social authentication method. We used the open source code of Signal to implement the three different methods of authentication. Our study materials are provided in a supplement and online at [redacted].

We recruited 21 pairs of participants (42 total participants) on campus, from July 17, 2018 to September 12, 2018. Participants were required to come in pairs, own an Android smartphone, and not have used Signal previously.

Participants were each compensated with $15 cash. Unlike previous work [14], we avoided giving an additional bonus to encourage participants to be security minded, to try to make the scenario more realistic. We did not provide participants with any instructions on the necessity of performing the authentication (in contrast to [13]), nor did we give them instructions on how to find or complete the authentication ceremony. We avoided giving any time pressure on participants to complete the task.

When participants arrived for their appointment, we presented them with the requisite forms for consent and compensation. We instructed them to download and install the customized Signal application being tested. We then read them a brief introduction describing the study conditions and their rights as study participants. We informed them that they would be placed in separate rooms. We also informed participants that a study coordinator would be with them at all times and would answer any questions they might have. We led the participants to their respective rooms, initiated video (and audio) recording, and instructed them to begin the survey. Throughout the study, coordinators were available to answer general questions but were careful not to provide any instructions that would aid in the use of the applications.

In the study, participants are asked to play the role of a parent and an adult child having a conversation. We asked participants to complete the scenario in two steps. In the first step, both participants were asked to have a normal conversation over Signal. In the second step, we asked one participant, in the role of the child, to ask for help filing tax forms, and to transmit a fake W-2 tax form to the other participant, who acted as the parent. We reminded the participants to treat the tax information as if it was their own. Despite the difference in roles, our intention was for both participants to complete the authentication ceremony. Participants were instructed to "talk aloud" as they performed the task, explaining their observations, actions, and reasoning.

We observed whether the participants used the authentication ceremony either in the first step or the second step, and which method the participants selected. We recorded statistics about their interactions with the ceremony, such as the time required to find it and time to complete it. During the task, participants could choose any of the three authentication methods. If participants transmitted sensitive data without completing the authentication ceremony, we noted this failure, and the study coordinators helped them initiate the authentication ceremony, using a method chosen by the participants. In all cases, once the task was completed, including using the authentication ceremony, participants were then asked to fill out a questionnaire. Following this, we asked them to repeat the process using each of the other two authentication methods. Before participants tried the other authentication methods, the study coordinator reset the verification status of their contacts. After using all three methods, participants answered additional questions on the questionnaire to compare the three authentication methods.

We did not control for the order in which participants tried each authentication ceremony method. Using techniques to control for order (e.g. randomization, or a Latin Square) would have required us to lead users to the authentication ceremony and tell them which method to use first, second, and third. We preferred a more natural scenario in which users made their own choices, thus revealing their preferences and avoiding any leading information about how to find the ceremony.

We recorded video of participants' phone screens during the study to ensure we could accurately measure which methods participants used and the time taken for tasks.

Mechanical Turk Survey

To generalize the results from our lab user study to a larger population, we conducted an IRB-approved, web-based survey asking respondents about their opinions on using social media for the authentication ceremony within secure messaging applications. We distributed the survey using the Qualtrics platform and recruited participants via Amazon Mechanical Turk.

To get reliable (non-spam) responses, we set the qualification requirement to a 96% acceptance rate. We limited our survey to the U.S. population, however our respondents are not necessarily a representative sample of the U.S. We initially collected 450 responses. We then removed any responses that were completed in less than 200 seconds, leaving us with 421 responses. The average response time for these participants was 10.70 minutes.

The questionnaire contained 42 questions, the majority of which were multiple-choice and Likert-type questions, with a few open-response questions. At the start of the survey, we provided an introduction about the purpose of the study and an implied consent form following the guidance from our IRB. The survey was divided roughly into the following groups of questions: (a) demographics, (b) privacy preferences, (c) usage of secure messaging applications, (d) usage of social media, and frequency of the interactions with phone contacts on social media (d) introduction to the importance of authentication in secure messaging applications, (e) instruction on how each of the authentication methods function, and (f) opinions regarding features that participants liked or disliked and their concerns about authentication via social media in open response questions.

Coding Qualitative Data

To analyze the data for open-response questions in the survey and interviews, three of the authors coded the data together using conventional content analysis. Any disagreements were resolved via discussion. First, we reviewed qualitative comments phrase-by-phrase and word-by-word to assign codes that classified users' comments with regards to a particular topic. Then, we used the constant comparative method to group codes into concepts and organized related categories by merging related codes and extracted themes.

Limitations

Due to our method of recruitment for our lab user study, our participants were primarily students and their acquaintances, and thus exhibited some degree of homogeneity.

For our app to work without approval from Instagram, we had to invite participants to our developer sandbox prior to the study. This may have predisposed them to use Instagram, though only a few did so.

Table 1: Order of trying each method of authentication

Method of Authentication	First	Second	Third
Social Media	12	21	9
In Person	10	12	20
Phone Call	20	9	13

Users needed to log into their social media accounts, even if already logged in with a separate application. Many had forgotten their password and had to reset it. This may have led to lower usability scores for social authentication.

Most of the time, the first phone call made with Signal did not go through properly, which has been reported as a bug in the Signal app (not our code). This may have led to lower usability scores for the phone call authentication method.

5 LAB USER STUDY RESULTS

In this section, we report on the quantitative and qualitative results of the lab user study. Our complete data set is at [redacted].

Demographics

We report on data collected from 42 participants who were recruited on campus. The participants are primarily young, with most (85.7%, N=36) between the ages of 18–24 and the rest (14.2%, N=6) between 25–34. They skew somewhat male (57.1% N=24) and the rest identify as female (42.8%, N=18). We asked participants to report their highest level of education, and most reported some college (66.6%, N=28), and the rest having a high school degree (16.6%, N=7), bachelor's degree (11.9%, N=5), or an associate's degree (4.7%, N=2). The majority of the participants (66.6%, N=28) self-reported their technical expertise as a beginner, with most of the rest rating themselves as intermediate (26.1%, N=11). Two participants whose majors were computer science and information technology rated themselves as experts.

Task Completion

We reviewed recordings of the study to determine which participants performed the authentication ceremony before sending or receiving the tax information. Table 1 shows the order in which participants tried the various methods. Two thirds of participants (66.6%, N=28) succeeded in performing the authentication ceremony, with about a quarter (23.8%, N=10) who did not even attempt it. Some participants (9.5% ,N=4) noticed the red bar that prompted participants to use the ceremony, clicked on it, and chose the phone call method, but they simply clicked the toggle to mark the conversation as verified without actually checking the safety number.

Time to Authenticate

Prior work has demonstrated that both QR code and phone call authentication can be completed in several minutes [14]. To test whether social authentication can be competitive with respect to time spent, we measured both the time it took for participants to register their social media accounts with Signal and the time to use the social authentication method in the authentication ceremony.

It took participants an average of 2:32 (minutes:seconds) to register their social media accounts with Signal. Only 10 participants authorized two social media accounts while the rest authorized just one account. Nearly all of the participants (N=39) authorized *Facebook*.

It took participants 00:34 (minutes:seconds) to use the method, as measured from the time they started the authentication ceremony to the time they clicked the toggle to mark the conversation as verified. Most of the participants (N=39) clicked on the profile picture of their conversation partner, and they spent an average of 00:18 (minutes:seconds) exploring their accounts. Most participants indicated that they recognized their friends just from their profile pictures. Those who did not tap on profile pictures mentioned that they didn't know this feature existed.

To check for order effects, we separated the timing data for each system based on which order the users tried each system. According to the Shapiro-Wilk test, the data is not normally distributed ($p < 0.05$). We next ran the Kruskal-Wallis test, which retained the null hypothesis for exploring social media accounts ($p = 0.422$) and rejected the null hypothesis for the time to use social authentication ($F(2, 42) = 9.885, p = 0.007$). A post hoc Tukey test showed a significant difference for time using social authentication when users tried it first (average 00:12) as compared to second (average 00:43, $p = 0.048$) and when they tried it first as compared to last (average 00:38, $p = 0.009$).

Comparison of Methods

Immediately after they tried each method, we asked participants to evaluate the usability of the method using a 5-point Likert scale ranging from *Extremely easy* to *Extremely difficult*. Figure 3 shows the scores for each method, with social media ranked as the most usable, followed by in person and phone call. The mean score is 1.64 for in person, 2.38 for phone call, and 1.61 for social media. An analysis of variance (ANOVA) on these scores yielded significant variation among the methods ($F(2, 123) = 7.488, p = 0.001, \eta^2 = 0.109$). A post hoc Tukey test showed a significant difference between the phone call and in person (p = 0.04) and between phone call and social media (p = 0.03).

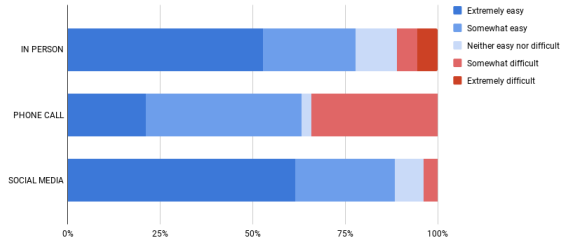


Figure 3: Single Ease Question (SEQ) scores for each authentication method (user study)

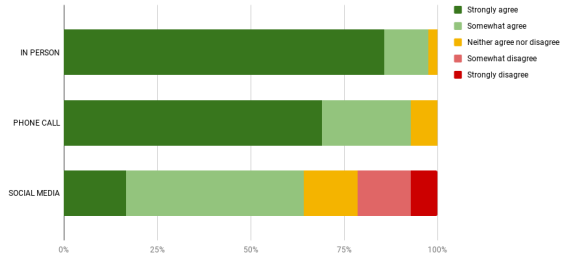


Figure 4: Trust scores for each authentication method (user study)

A one way Anova test showed that there was no significant difference in the usability scores of each system based on the order used (p = 0.449).

After participants had tried all of the authentication methods, we asked participants which one was their favorite. Social media was chosen the most (39.0%, N=16), followed by In Person (34.1%, N=14) and Phone call (26.8%, N=11). One participant did not chose any preferred method because they believed that all of the methods have major flaws.

After they had tried all of the authentication methods, we also asked participants to evaluate their agreement with the statement: *I trust this method of verifying the safety numbers in Signal.* The scores for each method are shown in Figure 4. The average rating for in person was 4.82, for phone call 4.64, and for social media 3.71. An analysis of variance (ANOVA) on these scores yielded significant variation among the methods ($F(2, 249) = 40.0, p < 0.001, \eta^2 = 0.243$). A post hoc Tukey test showed a significant difference between social media and in person (p < 0.001) and between social media and phone call (p < 0.001).

Because questions on trust and the favorite system were asked after the users tried all three systems, we assume that the order in which the systems were tried would not have a significant effect.

Participant Perceptions

After each authentication method they tried, we asked participants what they liked and disliked about that method in separate open-response questions. After coding, we identified the following themes:

Social authentication. Many participants liked that this method was fast and easy. Some liked that it worked asynchronously, meaning they could complete the ceremony independently without needing the other person to be simultaneously available. For example, P30 stated:

> *"I liked that I didn't even have to bother the other person in order to verify the privacy of the conversation."*

Other participants mentioned security, being automatic, working remotely (not having to be physically in the same location), and being able to use multiple social media accounts. P41 stated:

> *" I feel like most people you talk to, you are familiar with their social media accounts. So it would be effective enough in identifying fakes."*

The primary negative impressions were a worry that social media accounts could be easily hacked and that the method was not very intuitive. Some of these fears were misplaced, due to a misunderstanding of how social authentication worked. For example, one participant stated:

> *"If your social media account is compromised it could leak your key to someone else."*

Other concerns were that it was complicated, did not include mutual authentication, general distrust of social media, didn't work with people they didn't know on social media, was not secure, and was error prone (due to human judgment about authenticity of accounts). For example, P29 stated:

> *"It requires human judgment to make sure the profile pictures match, so that adds a little bit of uncertainty."*

We separately asked whether participants had any concerns about using social media accounts to provide the social authentication feature. In addition to worries expressed above, some participants did not trust the security associated with social media accounts and were concerned about a domino effect where if one gets hacked, all others are compromised. Another major point of concern was identifying contacts through pictures on their social media. Participants expressed concern for impersonator accounts that fake their friends' accounts. Other participants didn't like the idea of mixing security with social media, either because it was too public or because they didn't trust the platform. Many participants flat out stated that they found this method insecure.

QR code. Many participants liked that this method was easy, fast, and secure. Other positive themes mentioned were that it provided mutual authentication, was easy to understand, was reliable, and automatic. Participant 17 mentioned:

> *"It seems to be very reassuring because you know for a fact your dealing with your friend as you're face to face"*

The QR code method is unique in that it includes strong positive reinforcement – a physical action leads to a green check mark and a statement that the conversation is now secure, whereas other methods require the user to toggle a switch once they have verified the social media accounts or safety numbers. This positive feedback seemed important. Participant P33 stated:

> *"The literal scanning of a QR code is oddly gratifying."*

Another participant, P41, liked the idea of using this for new contacts:

> *"It was super quick and easy. And I feel like it would be better when you meet up with someone you don't know, to verify that they are who they say they are. I feel like it would make me feel a lot safer meeting someone new with this to verify."*

Ironically, the fact that you have to be in person was the leading cause of disliking this method. Many people don't want to meet up in person to verify a safety number. For some of these participants, it's nearly impossible because their friends live in other states or even outside of their home country. Other negative impressions were that it lacked clear instructions, was slow, and had a bug (didn't work the first time).

Phone call. Participants liked that they could verify the other person by voice (meaning they could recognize the voice of their friend), and that it was easy. Participant P5 stated:

> *"I liked that I could talk to him, so that I could be sure it was really him I was talking to."*

Other positive impressions were that it was fast, secure, had the safety numbers visible during the phone call, could work when participants were in different locations, had clear instructions, and the phone call was integrated. Three participants mistakenly thought the method was asynchronous because they toggled the verified switch and thought this activated a security feature. The primary negative impressions were the length of the safety numbers and a bug that prevented the first phone call from being completed for many users. The tediousness of the process could lead to unsafe assumptions. For example, P10 stated:

"I need to listen carefully if I really care about this. But I guess I can just assume that the 2 numbers are the same"

Other negative impressions were fears that it could be hacked (or someone could eavesdrop), that it was complicated, was not asynchronous, lacked clear instructions, was slow, and was not secure. One participant mentioned that if some one can impersonate his voice, then they can also modify the safety numbers and fool him. While this is not an easy attack, significant progress has been made recently [8].

Participant Understanding of Privacy

Immediately after completing the task, we asked participants whether they believed their conversation was private, and why or why not. Participants were split, with about half believing their conversation was private (47.0%, N=20) and about half saying they were not sure (47.0%, N=20), with just two participant believing their conversation was not private.

We coded the responses to this question and identified the following themes as reasons why participants believed their conversation was private:

- **Authentication:** Participants mentioned verification using a phone call or the QR code methods. When they mentioned the QR code method, they emphasized the fact that they physically met their conversation partners, rather than verifying the safety numbers.
- **Trusting the application:** Users mentioned Signal's professional UI, and others mentioned the notification in the application that their conversation was private. This trust was independent of whether users successfully completed the authentication ceremony or simply toggled the verified switch.
- **Low possibility of attack:** Some considered the probability of attack to be low, based on previous experiences.
- **Self filtering:** Some participants covered the (fake) social security number when they sent the tax form, or set auto delete feature on the message.
- **Signal contact verification:** A few participants considered the registration process, where Signal verifies their phone number, as demonstrating the application provided private conversations.

We also identified the following themes as reasons why participants believed their conversation was *not* private:

- **Unsure:** Some users said they did not have enough information to judge or were not sure if they were careful enough or had missed some steps. They were particularly unsure about the privacy of the tax information they had sent.
- **Lack of indicator:** A couple of users were suspicious about the privacy of the conversation because they didn't see an indicator for it.
- **Lack of transparency:** Some others blamed the application for lack of transparency, since it did not explain what it meant for a conversation to be private.
- **Possibility of attack:** Participants were concerned about possible attacks, especially impersonation and physical attacks on the recipients' side. One person worried that the tax information could be seen by anyone who had access to the recipient's phone.

Participant Understanding of the Ceremony

After participants had tried all three authentication methods, we asked them: *Why do you think Signal asks you to verify safety numbers?*. Participant responses had a high degree of variability, including mentioning security generally, mentioning privacy generally, believing it had to do with verifying the recipient's identity, preventing a spoofing (man-in-the-middle) attack, or expressing confusion. Most participants did not have a clear understanding of what the ceremony accomplished, but they often perceived that the safety number would somehow help enhance the privacy of the conversation, either by ensuring them that attacks are not happening or helping them verify they were talking to the right person. They generally did not make a connection between the safety numbers and encryption, nor did they have a complete picture of the threats they might face.

6 MECHANICAL TURK SURVEY RESULTS

In this section, we report the results of the Mechanical Turk survey. We discuss the feasibility of social authentication and how well the results of the user study generalize to a larger population. Our complete data set is at [redacted].

Demographics

All of our participants are over 18. About half (52.7%, N=22) of our participants are ages 25–34, and about a quarter (25.8%, N=109) are ages 34–44. Participants skewed male (60%, N=252), with fewer female (40%, N=168). Participants had a variety of education: less than 4 years of college (46.7%, N=197), 4-year degree (37%, N=156), Master's Degree (15.2%, N=64), Doctorate (0.9%, N=4). Participants self-reported their technical expertise with a mix of beginner (4.7%, N=20), between beginner and intermediate (9%, N=39), intermediate (42.2%, N=178), between intermediate and expert (31.1%, N=131), and expert (12.5%, N=53).

Feasibility

To determine the feasibility of using social authentication, we asked participants: *Now we want you to think about all*

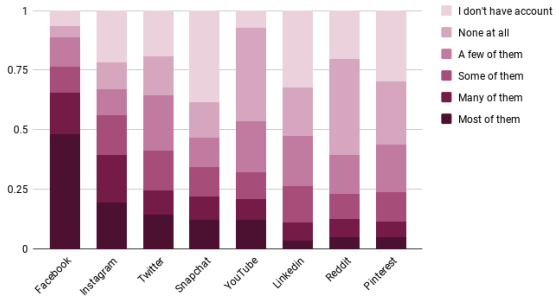


Figure 5: *How many of these people are you also friends with on the following social networks?*

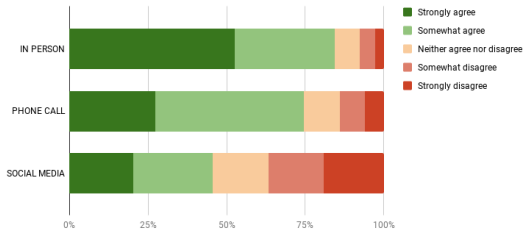


Figure 6: Trust scores for each authentication method (MTurk survey)

the people you normally communicate with using text messaging or instant messaging. How many of these people are you also friends with on the following social networks? We listed the same sites mentioned above. Figure 5 shows the results, demonstrating that social authentication will be feasible for at least some contacts for a quarter of participants on most social networks, with Facebook being the most widely used.

Comparison of Methods

We instructed participants about secure messaging applications, the purpose of the authentication ceremony, and comparing safety numbers. We then showed participants screenshots of each of the authentication methods and, after each method, asked them to rate how much they trust each method, using a Likert scale to agree or disagree with the statement: *I trust the following methods to enhance the privacy of my conversations.* Responses are shown in Figure 6. The average rating for in person was 4.26, for phone call 3.82, and for social media 3.10. An analysis of variance (ANOVA) on these scores yielded significant variation among the methods ($F(2, 1259) = 103.768, p < 0.001$). A post hoc Tukey test showed a significant difference between all pairs ($p < 0.001$ for each pair).

Comparing the trust scores, two points stand out. First, social authentication is least trusted and in person is most trusted, indicating that trust is much easier when mediated by physical confirmation. Second, there is a some amount of distrust for all methods, indicating that trust is difficult goal to attain for secure messaging applications.

Similar to the user study, we asked participants: *Which one of the above methods of authentication do you like?* The most preferred method was Social Media (37.7%, N=159), matching our user study, with Phone Call (34.4%, N=145) and QR Code (27.7%, N=117) ranking equally.

Participants were also asked to explain what they liked and disliked about each method, and their concerns about the Social Media method. After coding their responses, we found most of the themes from these open responses were similar to themes found from the user study. The primary differences were themes found in the user study due to experience with our particular implementation of social authentication. These included positive sentiment about not needing to leave the app to check profiles, and negative sentiments about a lack of clear instruction, not knowing what was being shared on their social media accounts, and the authentication not being mutual. Participants in the survey shared positive sentiments about the modern design and liked that they could verify multiple people in one sitting. They also noted concerns about the reliability of profile pictures.

7 DISCUSSION

Feasibility of Social Authentication

One of our primary purposes of this work was to test the feasibility of social authentication as a way to partially automate the authentication ceremony. Our user study demonstrated that participants find the method usable and fast, but they clearly did not spend much time verifying whether their contact's social media accounts had been faked or compromised. It would be even more difficult for users to judge the authenticity of a user's profile that they don't interact with often. In addition, the security of this approach improves as a user approves more social media accounts, but users authorized usually one and at most two accounts. This was a consequence of them not interacting with their friend on many accounts at once. Thus, even though an attacker would need to impersonate someone on all their social media accounts to compromise social authentication, it is not clear that users would use more than a few and it is not clear they would spot a clever attack.

Even if we could design a social authentication system that overcomes these flaws, based on participant feedback it would be very difficult to help users understand they are somehow improving their privacy by allowing Signal to access their social media accounts. Posting something on social media, even if users could understand it was just a public key, seems to them to be the opposite of private.

Trade-offs Among Methods

Participants identified clear trade-offs with the methods we presented to them. They find the in-person method trustworthy and understandable, but inconvenient due to needing to be in the same location. They find the phone call method convenient, but not very usable due to the long safety numbers that must be compared. They find social authentication works well remotely, automates the comparison, and can also work asynchronously, but they don't trust the use of social media accounts for this purpose. Ideally, users would like a method that is **asynchronous**, **automated**, **works remotely**, **is understandable**, and **trustworthy**.

Part of designing an understandable method is providing clear feedback to users regarding their actions and what they are accomplishing. For example, one thing users liked about the QR code method is that once they scanned the code, they viewed a large green checkmark and a message indicating their conversation was secure. While users could use more help understanding *why* scanning the QR code protected them, this feedback was helpful. With the other methods, users took some action (matching the safety numbers, checking social media profiles), but then had to switch the *Verified* toggle themselves. There was no clear connection between what they did and how this protected them.

Recommendations

Social authentication meets many of the participants' expectations for an authentication ceremony – it is asynchronous, works remotely, and is partly automated. One way to better meet expectations is to fully automate the ceremony using service providers that are more trusted than social media companies. A Signal user could register with additional providers, who each verify their phone number and publish their public key. The Signal app could then retrieve keys from these providers and compare them with the key published by the Signal server. Several additional points would need to be fleshed out with this architecture. For example, a method is needed for mapping users to providers, and a method for auditing providers for equivocation, such as CONIKS [6] would be helpful. Users would need help choosing providers and would still need help recovering from an attack if one was detected. One potential advantage of this architecture is that it could lead to interoperability among secure messaging apps, a weakness in existing applications.

8 CONCLUSION

We have integrated social authentication into Signal, as the first attempt to partially automate the authentication ceremony and translate it into a task that users can better understand. Our results are mixed. Participants find social authentication to be usable and fast, but they have significant doubts about its trustworthiness. Participants also spent too little time examining social media profiles, which would likely lead them to be susceptible to fraudulent profiles if an attack was occurring. However, our comparison of social authentication with other common methods helped identify principles for an improved authentication ceremony and clear trade-offs in how existing methods meet some principles but fall short in others. Our work illustrates that there is promise in automating the authentication ceremony and points the way for future work in this area.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers and our shepherd for their helpful feedback. This material is based upon work supported by the National Science Foundation under Grant No. CNS-1528022.

REFERENCES

[1] Hala Assal, Stephanie Hurtado, Ahsan Imran, and Sonia Chiasson. 2015. What's the deal with privacy apps?: A comprehensive exploration of user perception and usability. In *International Conference on Mobile and Ubiquitous Multimedia (MUM)*. ACM.

[2] Erinn Atwater, Cecylia Bocovich, Urs Hengartner, Ed Lank, and Ian Goldberg. 2015. Leading Johnny to Water: Designing for Usability and Trust. In *Symposium on Usable Privacy and Security (SOUPS)*. USENIX Association, Montreal, Canada, 69–88.

[3] Sergej Dechand, Dominik Schürmann, TU IBR, Karoline Busse, Yasemin Acar, Sascha Fahl, and Matthew Smith. 2016. An Empirical Study of Textual Key-Fingerprint Representations. In *USENIX Security Symposium*. USENIX Association.

[4] Amir Herzberg and Hemi Leibowitz. 2016. Can Johnny Finally Encrypt? Evaluating E2E-Encryption in Popular IM Applications. In *Workshop on Socio-Technical Aspects in Security and Trust (STAST)*. Los Angeles, California, USA.

[5] Ada Lerner, Eric Zeng, and Franziska Roesner. 2017. Confidante: Usable encrypted email: A case study with lawyers and journalists. In *Security and Privacy (EuroS&P), 2017 IEEE European Symposium on*. IEEE, 385–400.

[6] Marcela S Melara, Aaron Blankstein, Joseph Bonneau, Edward W Felten, and Michael J Freedman. 2015. CONIKS: Bringing Key Transparency to End Users.. In *USENIX Security Symposium*. USENIX Association, 383–398.

[7] Svenja Schröder, Markus Huber, David Wind, and Christoph Rottermanner. 2016. When SIGNAL hits the Fan: On the Usability and Security of State-of-the-Art Secure Mobile Messaging. In *European Workshop on Usable Security (EuroUSEC)*.

[8] Maliheh Shirvanian and Nitesh Saxena. 2015. On the security and usability of crypto phones. In *Proceedings of the 31st Annual Computer Security Applications Conference*. ACM, 21–30.

[9] Maliheh Shirvanian, Nitesh Saxena, and Jesvin James George. 2017. On the Pitfalls of End-to-End Encrypted Communications: A Study of Remote Key-Fingerprint Verification. In *Annual Computer Security Applications Conference (ACSAC)*. ACM, 499–511.

[10] Maliheh Shirvanian, Nitesh Saxena, and Dibya Mukhopadhyay. 2018. Short voice imitation man-in-the-middle attacks on Crypto Phones: Defeating humans and machines. *Journal of Computer Security* Preprint (2018), 1–23.

[11] Joshua Tan, Lujo Bauer, Joseph Bonneau, Lorrie Faith Cranor, Jeremy Thomas, and Blase Ur. 2017. Can Unicorns Help Users Compare Crypto Key Fingerprints?. In *Conference on Human Factors and Computing Systems (CHI)*. ACM, 3787–3798.

[12] Elham Vaziripour, Justin Wu, Reza Farahbakhsh, Kent Seamons, Mark O'Neill, and Daniel Zappala. 2018. A Survey of the Privacy Preferences and Practices of Iranian Users of Telegram. In *Workshop on Usable Security (USEC)*.

[13] Elham Vaziripour, Justin Wu, Mark O'Neill, Ray Clinton, Jordan Whitehead, Scott Heidbrink, Kent Seamons, and Daniel Zappala. 2017. Is that you, Alice? A Usability Study of the Authentication Ceremony of Secure Messaging Applications. In *Symposium on Usable Privacy and Security (SOUPS)*.

[14] Elham Vaziripour, Justin Wu, Mark O'Neill, Daniel Metro, Josh Cockrell, Timothy Moffett, Jordan Whitehead, Nick Bonner, Kent Seamons, and Daniel Zappala. 2018. Action Needed! Helping Users Find and Complete the Authentication Ceremony in Signal. In *Symposium on Usable Privacy and Security (SOUPS)*.

[15] Justin Wu and Daniel Zappala. 2018. When is a Tree Really a Truck? Exploring Mental Models of Encryption. In *Symposium on Usable Privacy and Security (SOUPS)*. USENIX Association.