Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study

Wei Bai, Michael Pearson, Patrick Gage Kelley*, and Michelle L. Mazurek
*University of Maryland, *Google*

***Abstract*—End-to-end encryption (E2EE) can effectively protect the privacy of online communication and has been adopted by several popular messaging apps. However, prior research indicates that many users have difficulty using E2EE tools correctly and confidently, as well as recognizing their security benefits, in part because of incorrect mental models. This paper takes a first step toward providing high-level, roughly correct information about end-to-end encryption to non-experts. In a lab study, participants (n=25) were asked about their understanding of E2EE before and after a tutorial we created, as well as which information they found most useful and surprising. Overall, participants' understanding of the benefits and limitations of E2EE improved. They found information about confidentiality, risks and weaknesses most useful, surprising, and compelling to pass on to others. Some confusion about integrity and authenticity remained. The results provide insight into how to structure new educational materials for E2EE.**

1. Introduction

End-to-end encryption (E2EE) is the best-known way to protect users' digital communications, as it prevents service providers as well as unassociated third parties from reading messages. In recent years, several popular messaging apps have adopted end-to-end encryption, either by default (WhatsApp, iMessage [1], [2]) or as an optional feature (Facebook Messenger, Telegram [3], [4]). As a result, after decades of use only in niche applications and communities, E2EE is now readily available and used by millions or even billions of users.

With this increased adoption, researchers have investigated how non-experts are using E2EE today. They have found many users do not understand the benefits and limitations that E2EE affords, and as a result may not use E2EE in the most secure fashion [5]–[10]. These difficulties are not only caused by poor usability or interface design; rather, they occur, in large part, because users hold what security experts would think of as "incorrect" mental models of E2EE. While users are not *wrong* to hold these mental models, this can lead to decisions or patterns of use that may put their security at greater risk than they realize. For example, users may wrongly believe other communications approaches (such as standard text messaging) are more secure than E2EE messaging, or may believe all secure systems are inevitably futile in the face of skilled adversaries, reducing the use of E2EE where it might provide important protections [6], [10]. Additionally, users may underestimate valid risks, such as vulnerability to malware at endpoints.

Ideally we would design systems where "users should not have to care about security" [10], and/or that — in interface if not in underlying technical operation — would "more closely align with users' existing mental models" [8]. However, both are likely to prove difficult for E2EE systems (at least in the near future), as existing models are quite far from the inherent technical characteristics of encryption systems, leading to significant misalignment. Instead, enabling users to make appropriate decisions about how to meet their privacy and security needs may require at least somewhat shifting mental models, perhaps via better explanations and education [8].

While it is neither possible nor desirable to ask all or even many users to become cryptography experts, we can and should work to improve mental models — particularly *functional* mental models that focus on how a technology is used rather than how it works [11]. This could help non-expert users understand basic threat models and mitigations, so that they can make informed and appropriate decisions about their own communication choices, such as what tools to use and how to use E2EE correctly within these tools [12].

Although some educational resources about E2EE exist [13], [14], they are often aimed at high-risk populations, such as journalists and activists. Also, users often need to take the initiative to visit these websites and access these resources. Our intent is different. We seek to provide non-high-risk, non-expert populations useful information in small doses, directly, while they are using E2EE tools. In this work we take an initial step toward this goal by exploring how to explain high-level E2EE concepts to non-experts, with a focus on what they find most important and surprising. In a qualitative lab study (n=25), participants were shown a short tutorial containing one or more modules explaining aspects of E2EE: a high-level overview; details about the kinds of surveillance risks E2EE can and cannot protect against; a debunking of common misconceptions; and a more detailed but only lightly technical description of how E2EE works. Before and after the tutorial, participants answered questions about their understanding of E2EE and its security properties. We also asked participants to provide feedback on the tutorial contents, to critique E2EE explanations drawn from popular messaging tools' current documentation, and to design a short explanation highlighting the most important aspects of E2EE.

Our goal was not to evaluate our exemplar tutorials as artifacts, but rather to investigate how users respond to different educational approaches, what aspects of E2EE they find most important or most surprising, and which elements they believe should be emphasized in future

educational interventions. While we do not expect many people will voluntarily complete even short formal tutorials like the ones we tested, we hope that insights from our study can be used to inform the design of a variety of naturally encountered educational efforts, such as interstitial informational screens within messaging apps, on-boarding flows, or help materials.

We find our tutorials effectively convey several high-level security properties of E2EE, including potential weaknesses at endpoints. They also correct existing misconceptions about who has access to messages in transit. Our risks module effectively conveys the relative risk of E2EE and non-E2EE communications, but our misconceptions module does not effectively address confusion about integrity and authenticity. As might be expected, the technical description we provided somewhat increased understanding, for some participants, but also added to misunderstanding for others, and was not considered particularly important or useful by our participants. The critique activities, meanwhile, revealed several points of confusion within existing messages. Overall, participants' responses suggest that emphasizing confidentiality, clearly conveying the limitations of E2EE protections, and reducing complexity are the most important properties for providing effective education in this space.

2. Related Work

2.1. Mental Models and Education for E2EE

Across many prior studies, mental models have been found to be fundamental in influencing users' decisions to (not) adopt encrypted communication tools. For example, users were reluctant to use encrypted email because they felt they "have nothing to hide" [15], and viewed routine encryption as paranoid and socially undesirable [16].

DeLuca et al. found users did not prioritize security and privacy when selecting messaging apps [5]. Relatedly, Abu-Salma et al. found "usability is not the primary obstacle to adoption" [6]; instead, fragmented user bases, lack of interoperability, and limited knowledge about security were significant barriers." The researchers also found that users held misconceptions about E2EE — e.g., believing that secure communications tools are inevitably futile in the face of powerful hackers — which hinder adoption. Wu et al. explored users' (mis)conceptions of encryption in depth, identifying four mental models [8]. Krombholz et al. identified key characteristics of encryption misconceptions via a series of drawing tasks [17].

Misconceptions found in these studies imply that improving mental models could support adoption of secure messaging tools. To date, little research has explored how to successfully do this. By comparing users' mental models before and after WhatsApp introduced E2EE, Dechand et al. found that WhatsApp's security info messages and generic media coverage were not effective [10]. Tong et al. tested a lock-and-key metaphor for public and private keys [18], with preliminary results indicating some improved understanding. We adopt this lock-and-key metaphor in a portion of our tutorial, described below. Demjaha et al. explored a variety of different short metaphors explaining the concept of E2EE but found that none were particularly successful [19]. Other work has included brief user-education materials as part of a larger study, but did not focus on how to design these materials effectively [7], [20]–[22]. There exist some educational resources about E2EE [13], [14], but they often target high-risk populations, such as journalists and activists; further, many users not in high-risk categories may lack the motivation to visit these websites and read the resources. Our work is among the first to explore educational materials for non-high-risk non-experts in a systematic way, and our qualitative approach allows us to explore how different aspects of an educational intervention may influence users' mental models.

2.2. Design Principles for Educational Material

We designed our tutorial using established educational principles and methods. Risk communication has been shown to be useful for conveying key points both in computer security generally [23]–[25] and for secure communication tools specifically [8], [12]. Therefore, we include risk communication as a module in our tutorial.

The principle of contiguity states that "the effectiveness of multimedia instruction increases when words and pictures are presented contiguously (rather than isolated from one another)" [26]. We designed our tutorial to include illustrative slides paired with narration and used pilot testing to refine the balance between visual content and written content. In line with recommendations to use first- and second-person point-of-view and make educational materials conversational [27], our examples of encrypted communication are framed to actively involve the participant as a sender/receiver.

Learning science also recommends giving learners opportunities to stop and think about what they have learned, as well as leveraging learning-by-doing ("knowledge and skills are acquired and strengthened through actual practice") [28], [29]. We incorporated these ideas by asking participants repeatedly to reflect on what they had learned in the tutorial, to consider what they would want to share with family and friends, and had them participate in a design task to exercise their new knowledge.

3. Methods

We designed an in-person study to explore the process of explaining E2EE to non-experts from several different angles. Participation took on average 63 minutes.

Goals and Non-Goals Our goal in this study was to explore the most important and useful information to convey to users when possible, as well as to obtain preliminary data about how best to convey that information. We hope the results can be useful in developing future educational materials and integrating them into apps and workflows, perhaps as short messages. As a first step, we developed an in-person, researcher-led tutorial as a *design probe*. Our goal was not to develop an optimal in-person tutorial, nor to evaluate our tutorial as an artifact in itself.

3.1. Study Procedure

We designed a qualitative study to investigate how participants' understanding could be influenced by intro-

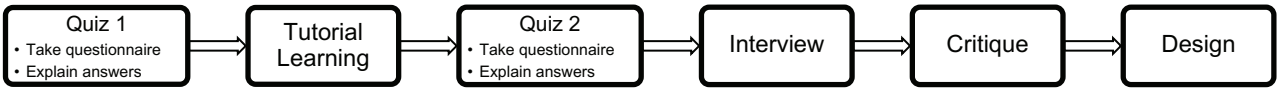

Figure 1. Study procedure flow chart.

ducing E2EE. Each participant was shown a short tutorial. Before and after the tutorial, participants were asked to answer a set of quiz questions regarding their perceptions of E2EE, then interviewed about their answers. We next asked them to comment on our tutorial, to critique some existing E2EE explanations from real apps, and to design a short message introducing E2EE in their own words. Figure 1 outlines our study procedure.

Initial Quiz After agreeing to our consent form, participants completed a closed-item quiz on paper to assess their understanding of E2EE. The quiz asked participants to imagine using a hypothetical E2EE messaging app, Textlight, to send a message to their friend Bob. We used a fictional app so that answers would not reflect participants' impressions of existing companies.

The questions explored participants' understanding of security properties of E2EE. To address confidentiality, we asked how difficult it would be for various adversaries to read the content of the message, including: hackers who could intercept communications, hackers who could infiltrate the app company, hackers who had previously installed malware on users' phones, ISPs, and government agencies. We also included two questions addressing integrity and authenticity, asking how difficult it would be for a third party to modify a message or send a response impersonating the recipient. All questions included five possible answers on a scale from "very easy" to "very difficult." Finally, we asked about the relative security of different communications tools, including SMS text messaging, mobile phone calls, email, and instant messaging apps with and without E2EE. After they completed the quiz, we asked participants to explain their reasoning for each answer. The quiz questions are listed in the appendix.

Tutorial and Second Quiz We then presented to participants our short tutorial, including narration and accompanying PowerPoint slides. After the tutorial, participants completed the same quiz questions again and were asked to explain why they had (not) changed their answers. We positioned this second quiz before any further study tasks so that answer changes would reflect the effect of the tutorial and not priming from other questions.

Tutorial Feedback We next asked participants to evaluate our tutorial, including how difficult it was to understand, how informative it was, and which parts they (dis)liked. We further asked participants to select the most and least surprising and important information, as well as the part of the tutorial they would most want their friends and family to learn about. These questions were designed in part to learn what did and did not work about our tutorial specifically. More importantly, however, this qualitative feedback was designed to allow us to develop hypotheses about how to best structure short messages that can be integrated into messaging tools without requiring users to seek out additional, extended training.

Critique Next, participants were shown two short texts (from six options) introducing E2EE. These were taken from app descriptions, official websites, and/or whitepapers of existing secure communication tools, including LINE [30], WhatsApp [1], Telegram [4], Viber [31], Threema [32], and ProtonMail [33]. We refer to these as Msg1–Msg6, in order. Each participant was presented only two to avoid fatigue. To avoid branding effects, we replaced the names of all six apps with Textlight. The example drawn from LINE (Msg1) follows; the other messages are provided in the appendix.

> *Letter Sealing is a feature that provides end-to-end encryption (E2EE) for chat room messages. E2EE is a communication system designed so that messages saved on our servers are encrypted and cannot be read by anyone except the sender and receiver of the message. Letter Sealing uses unique, user-specific encryption keys which allow users to safely and securely send messages to one another.*

We asked participants to give their opinions about both messages, mark phrases or sentences that explained E2EE clearly (or not), describe what else they might like these messages to tell them, and choose which message they preferred. This task had three goals: (a) to gain insight into perceptions of existing short educational messages; (b) to indirectly observe what they had retained from the tutorial and/or any remaining misconceptions, as revealed by their reactions to the messages, and (c) to provide some setup for the subsequent design task.

Design Finally, participants were asked to design a short message (200 words) about E2EE intended to teach others. Like the Critique task, this task had concurrent goals: to gain insight into how to design future short educational interventions, and to observe how participants were able to put knowledge gained from the tutorial into action. Participants' messages are included in the appendix.

3.2. Tutorial Design

One critical question concerns which pieces of knowledge are most important and useful to impart to users. While there are many potentially relevant aspects of E2EE we could try to explain [6]–[8], [10], [12], [18], [19], [22]–[25], we limited our investigation to only a few aspects for feasibility with our qualitative, exploratory method.

Specifically, we developed four modules: a basic overview, a module focused on risks E2EE can and cannot protect against, one explicitly debunking misconceptions reported in prior literature, and a more detailed (but not overly technical) description of how E2EE works cryptographically. The first three modules intend to cue functional mental models [11], which people could use to make decisions about how to protect themselves. The final module is intended to examine whether some very lightweight structural information [11] could help to support the functional models we primarily aim at. We

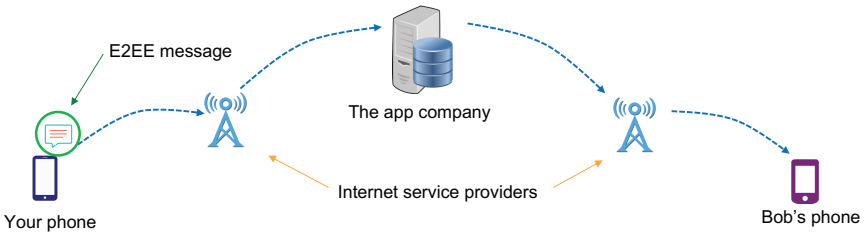

Figure 2. Example visualization used in the tutorial, depicting how an E2EE message is sent from "Your Phone" to "Bob's Phone".

mentioned E2EE could protect authenticity, but did not include how to perform authentication ceremonies, which is both platform-specific and in large part a usability issue. We leave whether and how improved mental models contribute to successfully performing ceremonies to future work. In order to explore the usefulness of the information contained in the different modules, we showed each participant one or more modules. All participants saw the basic overview, then zero to two other modules (see Section 4). Figure 2 shows an example from the tutorial.

Overview The basic overview briefly mentions confidentiality: "with E2EE, only you and your intended recipients can see the messages." We also briefly describe how E2EE protects integrity and authenticity, as well as endpoint weaknesses such as shoulder surfing, message forwarding, and malware on user phones.

Risks The risks module describes in detail risks to users when not using E2EE. We frame these risks by first illustrating, at a high level, how messages transit through ISPs and messaging app companies on their way to the recipient. This illustration aligns with those presented in prior work [7], [22], with the addition of ISPs. (As we will see below, the role of ISPs in the messaging process proved surprising to many participants.)

We first highlighted many possible risks of unprotected communication within this ecosystem: interference at endpoints, at the ISP, at the messaging app server, and in transit in between. We then discussed how point-to-point encryption (e.g., webmail over TLS) could mitigate some of these risks, before finally detailing the additional mitigations provided by E2EE.

Misconceptions The misconceptions module focused on three misconceptions or misunderstandings identified in prior work. First, prior research has identified a belief that encryption is "futile" because no privacy mechanism could be strong enough to protect from powerful attackers like governments or highly skilled hackers, especially if the mechanics of the encryption system are known to the attacker [6], [8], [10]. We presented this misconception and then briefly explained the proven strengths of (properly configured) encryption algorithms against even well resourced and knowledgeable attackers.

Second, many users believe that standard audio calls and SMS messages are more secure than E2EE messaging apps [5], [6], [10]. We clarified that E2EE messaging is the most secure of these options. Finally, some users mistakenly believe that authenticity and integrity are controlled entirely by username and password security [6]. We explained that even if usernames and passwords are not compromised, without encryption attackers may be able to tamper with messages or impersonate other users.

Crypto Knowledge Our final module explains public key encryption using a lock-and-key metaphor adapted from Tong et al. [18]. Participants were told that at app installation, a public lock and private key pair would be created on their phones. All users' public locks would be stored at the app company, referencing the key-directory model commonly used in messaging apps. To send an encrypted message to a friend, the participant's device would request the friend's public lock and use it to lock the message; only the friend, via possession of the matching private key, would be able to read the message. Prior work has obtained mixed results on the utility of explaining the underlying cryptography model [8], [18], [19]. We included this module to investigate both how it would affect users' overall understanding and whether they would find it interesting or important.

3.3. Recruitment

We recruited participants who were 18 or older, use messaging apps at least once a week, and did not possess significant prior cryptographic knowledge, as reported on a screening survey. We placed flyers around the University of Maryland campus and surrounding neighborhoods, sent emails to university distribution lists, and advertised on Craigslist for Washington, D.C. All interviews took place in person, on our campus, and all sessions were audio recorded (with permission). Participants were paid $20 for completing the one-hour study and reimbursed for parking when applicable. The study protocol was approved by our organization's ethics review board.

3.4. Data Analysis

Two researchers transcribed three sessions, and the remaining 22 were transcribed by an external transcription service. The majority of our collected data was qualitative. We applied textual microanalysis to analyze these data [34]. Two researchers iteratively developed a codebook using responses from three participants to capture key themes and ideas. The two researchers then independently coded all participants' responses. We use Krippendorff's α to measure the inter-coder reliability [35]. The average α across the codebook is 0.85, which is considered reliable [35]. After calculating reliability, the researchers met to resolve all disagreements.

Because of our qualitative focus and thus the limited sample size, we use changes in quiz responses descriptively, to contextualize participants' qualitative responses to the tutorial. We do not apply hypothesis testing or attempt to make statistical comparisons.

3.5. Limitations

Our sample is small, and our participants were younger and more educated than the general population, potentially limiting generalizability. Most participants (see below) had heard of encrypted messaging, but few actively used it, and all had limited knowledge about encryption and related technologies. We therefore believe our study can provide insight about moderately tech-savvy users who might consider adopting encrypted messaging. Results from this exploratory study can be used as a starting point for developing E2EE explanations and validating them with larger samples.

As in any interview study, participants may have answered quickly rather than thoroughly [36]; to mitigate this, we used a variety of tasks and follow-up questions. Demand characteristics suggest that participants might mirror the content of the tutorials back to the interviewer in an effort to tell us what they thought we wanted to hear [37]. In this study, attempts to mirror content demonstrate what the participant has (not) learned and therefore are not a major source of concern.

More broadly, our tutorial necessarily encodes some of the researchers' judgment about what E2EE concepts are most important for users to learn; this may affect participants' information retention and their responses about importance. We attempted to mitigate this by providing a broad range of content for participants to consider and by endeavoring to present all concepts as equally important.

Finally, our results, in which participants focused on the material for several minutes, may not translate directly to real-life behaviors like glancing at introductory material when using a messaging app. Nonetheless, we believe that insights from our study — particularly which concepts proved most important and surprising — provide a foundation for future research into shorter interventions that users encounter in passing.

4. Participants

In total, 62 people completed our screening survey. We recruited 26 participants on a rolling basis, eliminating one who withdrew mid-study. Demographics are shown in Table 1. Fourteen of 25 participants were female and 21 were between the ages of 18 and 29. The participants tended to have an educated background, with 19 holding a bachelor's or master's degree.

Only one participant reported an education in computer security, computer science, computer engineering, or IT, and none reported familiarity with cryptographic algorithms. Eighteen reported either not having heard of or not having used encrypted or secure messaging, while five reported that they actively use such messaging. However, two of these five used only non-E2EE WeChat.

Participants reported using many messaging tools at least once a week: Facebook Messenger (20), iMessage (16), Snapchat (12), SMS text messaging (12), Google Hangouts (8), and WhatsApp (6). Four or fewer reported using each of WeChat, Telegram, KakaoTalk, Google Allo, GroupMe, Instagram Messenger, and Twitter.

As mentioned in Section 3.2, we created four modules to test with our participants. For this exploratory, lab-based study, each participant started with the overview module. Some participants then immediately moved on to the secondary quiz, interview, and other tasks, while others saw up to two other modules. We randomly distributed participants among five combinations of modules we found interesting: 15 participants were shown the risks module, 10 saw misconceptions, and only 5 saw the cryptographic knowledge module, which we expected (based on prior work) would be least useful. We did not compare across participants who saw different combinations of the modules either qualitatively or statistically. Rather, we wanted to gain an initial indication of how the different explanation types worked both separately and in combination, which could help us gain qualitative insights about how participants interpreted each module.

ID	Gen.	Age	Eth.	Mod.	Crit.	Familiarity
P1	F	18-29	Asian	ORC	3,4	use frequently*
P2	F	30-39	Black	OR	5,6	not heard of
P3	F	18-29	Black	O	1,2	not heard of
P4	M	30-39	Black	OM	1,3	use frequently
P5	F	18-29	White	ORM	2,5	not heard of
P6	F	18-29	White	ORC	3,6	not heard of
P7	F	18-29	Asian	ORM	2,5	heard, not use
P8	F	18-29	Black	OM	1,5	heard, not use
P9	M	18-29	Asian	ORC	4,5	heard, not use
P10	M	18-29	Black	O	3,5	not heard of
P11	F	18-29	Black	OR	1,4	heard, not use
P12	M	50-59	N/A	ORM	1,6	heard, not use
P13	M	18-29	Asian	OM	2,4	use frequently
P14	F	18-29	Asian	OR	2,3	not heard of
P15	M	18-29	Hispanic	ORM	3,4	not heard of
P16	M	18-29	Black	ORC	1,2	use occasionally
P17	F	18-29	Other	OM	3,6	not heard of
P18	M	18-29	Asian	O	2,6	not heard of
P19	F	18-29	White	OR	4,5	heard, not use
P20	M	18-29	Asian	ORM	1,3	not heard of
P21	F	18-29	White	ORC	1,6	heard, not use
P22	F	18-29	Asian	OR	2,6	use occasionally
P23	M	18-29	White	OM	4,6	heard, not use
P24	M	18-29	Asian	O	1,5	use frequently*
P25	F	30-39	Black	O	4,6	use frequently

TABLE 1. PARTICIPANT DEMOGRAPHICS INCLUDING, GENDER, AGE, ETHNICITY, TUTORIAL MODULES, MESSAGES CRITIQUED, AND PRE-EXISTING FAMILIARITY WITH ENCRYPTED OR SECURE MESSAGING. TWO PARTICIPANTS WHO ONLY USE WECHAT BUT IDENTIFIED THEMSELVES AS FREQUENT USERS OF ENCRYPTED TOOLS ARE MARKED WITH *. MODULES INCLUDE OVERVIEW (O), RISKS (R), MISCONCEPTIONS (M) AND CRYPTO KNOWLEDGE (C).

5. Results

We will discuss what participants learned from our tutorial, which incorrect mental models persisted, and some other feedback.

5.1. Tutorials improved mental models

Participants' understanding of E2EE improved regardless of which tutorial modules they saw. This is reflected in changes to their quiz answers post-tutorial, as well as their interview comments.

5.1.1. Improved understanding of messages in transit. Our quiz included two adversaries related to threats in transit: eavesdropping hackers and ISPs. Participants learned that E2EE could protect their messages from these two adversaries: 13 and 10 participants lowered their scores (perceiving eavesdropping to be more difficult) in these two questions, for hackers and ISPs, respectively.

Before the tutorial, participants held several misconceptions about these potential adversaries. Seven thought these adversaries could not learn message content as they were not part of the conversation, and four (one overlapping) were surprised to learn that ISPs were involved in the communication. For example, P20 said "[ISPs] are even more outside, like, the actors involved in the communication, [they are] not the company making the app and they're not any of the users, it will be very hard for them to open up the message itself." P22 commented, "I thought this was one to one relations. Even though they provide the service. ... Like KakaoTalk, I believe that company can access to my messages, but I've never thought that the provider, Verizon [could]. I thought that they just provide the transportation." P21 did not realize the recipient's ISP was involved, believing the app company would send the message directly to its recipient.

On the other hand, six participants believed before the tutorial that ISPs could potentially learn any message, since they provide the internet service for phones. As P14 said, "I think they have access to everything you send through the network." After the tutorial, all six participants recognized that encryption (point-to-point or end-to-end) could address this issue.

5.1.2. Protection from app companies and governments. After the tutorial, participants understood better that E2EE could protect their messages from app companies and governments; 17 and 19 participants lowered their scores for these adversaries, respectively.

Before the tutorial, 17 participants who thought E2EE was vulnerable to these adversaries incorrectly mentioned plaintext messages stored at the app company, where governments could request them. In line with prior findings [6], [10], 11 suggested that these relatively powerful adversaries could simply break the encryption.

After the tutorial, 22 participants recognized that E2EE prevents plaintext messages from being stored at the app company and provides a defense, even against powerful adversaries. P14 stated, "Using E2EE there will be a shield. ... The company would know there is a message that exists, but they wouldn't know what content it contains." P15 similarly commented, "It's easy for the government to get a copy ... but with E2EE they won't be able to break the encryption to see what's inside."

5.1.3. Better recognition of E2EE limitations. After the tutorial, 10 participants rated the risk at the endpoints (even when using E2EE) higher than they had initially, and 21 participants correctly understood that E2EE could not protect against malware on their phones (up from 16). Before the tutorial, six participants thought encryption could protect against malware at endpoints: "Malware [is] getting access to the activities, but ... the message is still going to be encrypted" (P8). Two participants did not understand before the tutorial how malware related to message privacy at all; P3 said malware is not "part of the conversation." P8 updated his opinion after the tutorial: "If malware's already installed on your phone, then it's not really protected because they have access to the original message before it's even encrypted." P12 was very specific in pointing out that E2EE could protect against "two out of those three risks: the transit ones, and the app company, but not the user phones."

5.2. Remaining (and new) concerns

Although our tutorial improved participants' understanding overall, some participants retained existing misconceptions and some even developed new ones.

5.2.1. Confidentiality concerns remain. After the tutorial, eight participants remained unconvinced that E2EE could provide sufficient confidentiality from powerful adversaries. For example, P2 argued that government agencies have "both the resources and the knowledge base to crack an algorithm." Similarly, P6 continued to believe that the app company would inherently have enough knowledge about the encryption process to access messages: "If the company creates the encryption or whatever, I'm assuming there is some sort of maybe possibility of them decrypting it ... the company still has your lock, meaning even if it's encrypted wouldn't they be able to at some point decrypt it?" P21 suggested skilled hackers could similarly overcome E2EE: "Because I'm thinking of it literally ... I would imagine that if you work in a locksmith office, if you're an expert in keys and locks, you might not have somebody's key but you would be able to get into their house because you are an expert."

5.2.2. E2EE is less secure than other communication methods. Four participants ranked SMS, email, or a direct phone call as the most secure communications method even after the tutorial. Some argued, not unreasonably, that voice conversations are more transient than written messages: "It's just talk. That's most likely to be forgotten" (P10). P11, however, preferred phone calls as a means to establish authenticity: "If I know my friend or someone in my family, it's very hard to mimic that voice .... With a phone call, I'd be like, 'Oh, that's not you.' "

5.2.3. Point-to-point encryption is devalued. All modules emphasized the benefits of E2EE. The risks module explicitly distinguishes point-to-point encryption from E2EE, but also mentions that most non-E2EE messaging apps are encrypted point-to-point. This nuance was not entirely absorbed. P21 recalled that non-E2EE messaging apps "[take] out the first threat even though there are the other threats. So maybe it's not quite as easy" as with no protection; however, she still ranked the security of such apps behind standard voice calling. Among the 15 participants who viewed the risks module, 10 retained their security ranking for these apps, among whom seven still ranked non-E2EE messaging in fourth or last place. Four other participants reduced their ranking for these apps. This suggests researchers, designers, and advocates promoting E2EE should be careful not to inadvertently push consumers toward less-secure options when E2EE is unavailable or undesirable.

5.2.4. Integrity and authenticity improve but remain confusing. Our results align with prior findings that users do not effectively consider integrity and authenticity [6], [8]. Before the tutorial, most participants did not think integrity (14) or authenticity (18) could easily be violated in

E2EE communication; however, many did not relate this difficulty to E2EE, instead believing that such violations would be inherently difficult in any messaging system. Eleven participants believed these violations would require multiple unlikely steps, and three believed an attacker could never be fast enough to alter a message in transit. P19 said, "It would have to happen while the [sender]'s typing, cause like if they get the message instantaneously, then...They'd have only like a second to change it." Eleven participants, meanwhile, believed integrity and authenticity were determined entirely by protection of usernames and passwords, rather than any vulnerability in transit.

After the tutorial, participants were more likely to recognize that E2EE could protect integrity and authenticity. As P2 explained, messages "have the end-to-end encryption, and if the phone's not stolen, it's going to be difficult." While participants believed our assertion that E2EE can assure integrity and authenticity, exactly how E2EE provides this protection remained muddled. Seven still confused integrity with confidentiality. For example, P25 remarked that "there's really nothing for him (the attacker) to modify because he doesn't know the contents of the message." (To be fair, properly designed E2EE tools should protect both confidentiality and integrity without requiring the user to make a distinction.) Further, two participants continued to believe that a limited time window for action would prevent an integrity or authenticity violation, and other participants continued to conflate authenticity with username and password issues. P25 suggested that "E2EE protects against message modification and impersonation. Not even usernames and or passwords can be stolen or guessed."

Relatedly, the integrity and authenticity discussion in the misconceptions module was cited by several participants as the least effective tutorial section, with three participants explicitly calling it unclear or confusing. This may be because these concepts are inherently difficult to explain, because we explained them poorly, or both.

5.2.5. E2EE protects against malware. While many participants increased their understanding of endpoint threats, five participants indicated (in the post-tutorial quiz) an increased belief that E2EE could protect against malware on their phones. As P5 explained, "Even if you put the malware on the phone, it's not gonna give you access because the data is so well encrypted." P19 concurred that "By using E2EE, you decrease the likelihood that someone can hack into your phone." These results underscore the importance of ensuring that educational interventions do not give users a false sense of security.

5.3. Other feedback

Finally, we report additional feedback about our tutorial, as well as the existing texts participants critiqued.

5.3.1. The tutorial was generally well received. Overall, participants rated the tutorial as easy to understand (mean 4.9 on a five-point Likert scale) and informative (mean 4.5). While these scores may partially reflect demand effects, participants also gave thoughtful comments. P8 said "I think it was easy because you used very clear, simple, and concise language. . . . The visuals really, I think, facilitated my understanding"; others agreed. Some, such as P20, said they gained new knowledge: "I learned a lot that I didn't know before, clarifying how messaging apps actually work, the risks and how encryption can help protect you." Participants including P3 also said the length was about right: "Since it's a short presentation, I'm sure there was more to learn. However, I don't think anyone needed to know more than that." A few participants, however, identified tutorial elements that were not clear, primarily integrity and authenticity (discussed above).

5.3.2. Wording should be clear and simple. The critique task revealed several instances where overly complicated explanations, jargon, and related issues inhibited understanding. For example, three of eight participants who viewed Msg3 struggled with technical terms like ISP, network administrator, and third parties. Both Msg2 and Msg5 mention that E2EE is applied automatically. Participants preferred the shorter, clearer wording in Msg5 (five likes, no dislikes) to the longer version in Msg2 (two likes, three dislikes).

Three of eight participants who saw Msg2 were confused by the "them" in *"Many messaging apps only encrypt messages between you and them."* P18 correctly guessed "they" were the messaging apps, but P3 assumed that "they" were message receivers and consequently misunderstood the rest of the sentence.

6. What knowledge should we impart?

We discuss which information was most salient to our participants: what they found most important and surprising, and what they would choose to tell others.

6.1. Confidentiality is most significant

Unsurprisingly, confidentiality was most significant to our participants. It was selected 13 times as the most important information and nine times as the most surprising, more than any other theme. P3 said "the inability of other parties to read your messages" was most important because "that actually is the primary purpose" of E2EE. P15 was surprised that "the internet service provider and the app company . . . may still get a copy of the message, that is protected by this wall, that is nearly impossible to break. So they can see you sent a message, but they can't see what the message says." P14 was similarly surprised that with E2EE, even the government could not read message content: "I thought the government could always read what you have sent because, like, you always heard in the news that some student said something stupid on their phone and they got deported or something. . . . So I always thought that the government can read everything you send. But now today I realize, oh they can't."

Seventeen participants said they would tell others about confidentiality. P8 would convey that E2EE "is going to be the most secure way that you can send messages. . . . It has kind of like a very, very, strong protective layer around the message so that people can't intercept the message and retrieve its contents. It will only be accessible to the intended recipient."

Confidentiality also proved important during the critique and design activities. Almost everyone pointed positively to mentions of confidentiality in existing messages, and 20 participants mentioned it in their own messages. P25's message said, "E2EE is a great way to ensure that your messages will be securely transmitted to the recipient of the message originally intended for. ... No one has direct access to the text contents but the sender and the receiver of the information. Not even the monitoring company, hackers and the government can see the message."

6.2. Risks usefully differentiate E2EE

Of 15 participants who saw the risks module six reported that risks were most important to discuss and five reported risks as most surprising. These participants emphasized the importance of clearly differentiating E2EE from non-E2EE communications. For instance, P6 said "I think, again, knowing the risks of the non-E2EE and then really comparing it to how is this better. So that's really the most important." P11 said, "I think they just want the bottom line, like okay, what is the difference? So, I think the risk factors are probably the biggest takeaway."

When asked what they would tell friends or family about E2EE, 10 participants (seven who saw the risks module) specifically mentioned risks. As with importance, these participants used risks to explain why E2EE was better than non-E2EE: P5 said he would start by explaining the risks of non-E2EE messaging and then explaining that E2EE "really protects it [your message] while it's in transit and while it's at the app company. ... It can't be modified or impersonated or whatever. ... [E2EE apps] basically protect your message along the entire pathway and it's the most secure way to send a message." P19 agreed that "showing them the benefits of encryption so it's less likely your information will be hacked would probably be beneficial," particularly in the context of many data breaches in the news.

We saw similar themes in participants' designed messages. P2 wrote, "Your information is exposed every time you utilize your mobile device. There are three potential exposures, the service providers, the app company, and hackers. ... To protect yourself against these parties, please ensure your device has E2EE protection." P7 began with a detailed message about risks before explaining how E2EE could address these risks.

6.3. Explaining weaknesses can be important

We introduced three endpoint weaknesses in the tutorial — shoulder surfing, message forwarding, and malware — but malware primarily drew our participants' attention. Four and five participants reported that this information was most important and surprising, respectively. P6 explained that "If you're presenting this as a product, I think it's important to [be] realistic and inform people." P18 agreed: "The malware piece is important for me to know if I'm specifically looking at TextLight company and what is not protected ... because I want to know to what extent is this actually going to be secure?" Similarly, four participants said that if they were to introduce E2EE to other people, they would mention the limitations, and 13 participants acknowledged limitations we discussed straightforwardly in their designed messages.

Among the critiqued messages, only Msg3 pointed out possible weaknesses: *"But please remember that we cannot protect you from your own mother if she takes your unlocked phone without a passcode. Or from your IT-department if they access your computer at work. Or from any other people that get physical or root access to your phones or computers running TextLight."* This information was marked as especially useful by five participants out of eight who saw this message. These participants noted positively both the content of the information and the humorous tone. (However, two participants including P15 noted that "root access" was jargon: "I don't know what root access means. What happened?")

6.4. Cryptography details: Use with caution

We observed mixed reactions to technical details presented in our crypto knowledge module and in the existing messages participants critiqued.

Only one of the five participants who saw this module selected it as important, saying "you cannot tell people this thing is so good, but you don't tell them why" (P1). None mentioned crypto knowledge or the lock and key metaphor when asked what they would tell their friends and family about E2EE. P18, who did not see the module but did critique Msg2, which used the same metaphor, remarked that this information provided "a little bit more understanding of what encryption could look like, even though it's not probably how it works."

Some participants were reassured to learn that a secret key was known only to the user and their devices, as pointed out in Msg4 and Msg6. P12 noted that this "eliminates all possibility on service providers ... they don't have the encryption keys in order to break the codes." Eight participants' design-task messages mentioned cryptography, including three that mentioned keys known only to the user and not the app company. Interestingly, seven participants who didn't see the crypto module said they wanted to learn more about how E2EE works, and one who did wanted even more details.

Other participants were confused by technical details. "This [tutorial] is intended for new users, people who don't know anything, like me. This [crypto explanation] is helpful, but I don't need to know exactly how it happens" (P1). Other participants found specific details confusing: P23 associated a key with a password and asked, "How do you have a key that you have to put in every time to see a message?" P5 was confused by the assertion in Msg5 that *"the user is in control over the key exchange,"* asking, "Do I have to do something? How does that key exchange happen?" Two other participants expressed similar feelings.

Seven of eight participants who critiqued Msg2 liked the lock-and-key metaphor, and three thought it explained encryption well, but three found specific language confusing. P3 said that the term "special key" was "flowery" and "corny." P14 objected to "added protection": "What does it mean by 'added protection?' I don't get it."

Msg5 emphasizes that *"the end-to-end encryption layer passes through the server uninterrupted; the server cannot remove the inner encryption layer."* This explanation increased four of eight participants' confidence in

E2EE protection, but five worried it was too technical or even boring. P8 couldn't picture what this would look like: "I didn't understand if it was an inner layer or outer layer, whatever that meant. It was just frustrating."

Of eight participant-designed messages mentioning cryptography, four tried to use the lock-and-key metaphor, and three made minor technical errors, including suggesting that the lock and key themselves were encrypted or providing a symmetric-encryption-like explanation.

6.5. Other concepts were less important

Integrity and authenticity, comparison to other tools, and algorithm security received little attention.

Comparison with other tools Four of 10 participants who saw the cross-tool comparisons in the misconception module mentioned it as most important, and two as most surprising. P7 found most surprising"the misconception that E2EE is more secure than the other methods, 'cause I thought that phone calls was the most secure form since it's using your voice and not actually recorded." Only two participants included a comparison to other communications mechanisms in the messages they designed.

Algorithm security The misconceptions module also included information about algorithm strength, intended to disrupt the assumption that an expert in cryptography can always break it [6], [10]. Of 10 participants who saw this, only two mentioned it as important and two as surprising. P5 was surprised that "it would take like a million years or something . . . 'cause I feel like you always hear about hackers." Four of these 10 participants mentioned algorithm security in their designed messages as a way to emphasize confidentiality. P20 wrote that the algorithms "are extremely strong and would take third parties an impossible amount of time to crack and decrypt."

Integrity and authenticity Integrity and authenticity were discussed briefly in the overview module seen by all participants, then in more detail in the misconceptions module seen by 10 participants. Few selected these concepts as important (1) or surprising (2). Three participants mentioned integrity and/or authenticity when choosing information to tell others, and six included them in the messages they designed. This unpopularity may relate to the fact that we were relatively unsuccessful explaining these concepts to our participants (as described above).

7. Conclusion and Recommendations

Our study takes a first step in exploring how to improve non-experts' mental models of end-to-end encryption. We designed a tutorial with four modules: a brief overview, risks associated with non-E2EE communication, corrections to three common misconceptions, and a high-level description of how E2EE works. We evaluated what our participants learned, what misconceptions they retained, and which information they found most valuable.

Our results suggest that participants' overall understanding of E2EE improved after the tutorial. We observed improvement in understanding how E2EE can protect confidentiality, how it compares to non-E2EE mechanisms, and its limitations. Participants could better recognize potential adversaries, such as internet service providers, they had not previously considered; they also newly recognized that E2EE can provide protection even against powerful adversaries like app companies and governments. Further, the tutorial helped participants recognize that E2EE cannot protect against endpoint threats such as malware. We believe these concepts — appropriate mental models of risks, threats, and protections — are crucial to users' ability to make meaningful decisions about their communications. This study shows that educational interventions on these concepts can improve understanding.

Although our tutorial provided a net improvement, participants did maintain some existing concerns and misunderstandings, and a few developed new ones. Our attempt to clarify authenticity and integrity — in particular to disentangle it from password security — was not successful. Further, some maintained their belief that standard voice calls are inherently more secure than any text communication, showing our tutorial doesn't yet capture nuanced differences in threat models between mediums.

In practice, of course, most users will not have the time or interest to complete a tutorial like ours. We therefore distill — from our participants' responses to the tutorial, their critiques of existing messages, and their efforts to design new messages — key elements that we hope can translate to shorter and more varied educational interventions, such as those on apps' webpages and help documentation, in their installation descriptions and onboarding, or in interstitial screens.

Confidentiality is the primary ingredient. Our participants placed the most importance on conveying how E2EE can protect users' messages from being read by various adversaries. This can be supported by references to risk, algorithm strength, and the comparative security of other communications. Perhaps most important is to explicitly explain protection against powerful adversaries, which was most surprising to many participants.

Risks can be used to support comparison. One way to effectively emphasize confidentiality is to detail the risks that distinguish E2EE from, e.g., point-to-point encryption. This improves threat models and can increase the sense of security, potentially motivating adoption.

Technical details only in small doses. While some participants wanted to know how E2EE works, most did not find it critical, and we observed a strong risk of misunderstandings. Previous work noted that commonly used metaphors may fail because they are *structural* (how the system works) rather than *functional* (what it can do) [19]. We similarly observed that technical details were most effective when functional. In particular, emphasizing that the secret stays only on users' devices appears to improve perception of security without creating confusion.

Clarifying limitations is important. Participants found value in clarifying what E2EE cannot protect against. Commercial entities may hesitate to call attention to drawbacks of their products, but our evidence suggests it can improve mental models that might otherwise limit adoption. (Of course, this must be done carefully to avoid undermining the overall usefulness of the tool.)

Explaining integrity and authenticity may not be worth it. Our unsuccessful explanations of integrity

and authenticity, combined with prior related findings, suggest conveying these nuanced concepts effectively will continue to be challenging. Relatively few participants regarded these concepts as directly important; instead, most absorbed them into their model of confidentiality. (This aligns with the approach in [7] of redefining authenticity in terms of confidentiality.) As such, we recommend against trying to explain integrity and authenticity independently, without considerable investment in identifying and testing a more comprehensible path forward.

Strive for simplicity. Various comments from our critique session demonstrate that jargon not only obstructs users' understanding, but also annoys them. Participants generally frowned on overly complex explanations. All the texts we examined appear to have been written to be colloquial, but we suggest even further simplification.

We hope these recommendations can inform the design of interventions that can be integrated more naturally into users' communication workflows. In particular, they must be validated across a broader demographic range of participants and in real-world situations when learning about E2EE is not the user's primary task. Future work should also examine how to tie these high-level concepts more directly to concrete tasks such as choosing a communication medium, turning on E2EE mode (when not automated), performing an authentication ceremony, or noticing a key change. Overall, mental models that are better aligned with how technologies function will enable more effective and private use of communications tools.

Acknowledgements

This material is based in part upon work supported by the U.S. Air Force and DARPA under Contract FA8750-16-C-0022. Opinions, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the U.S Air Force and DARPA.

References

[1] F. Inc., "Whatsapp," https://www.whatsapp.com/, 2018.

[2] Apple, "iOS Security, iOS 11.0," Jan 2018.

[3] "Facebook Messenger," https://www.messenger.com/, 2018.

[4] "Telegram Mobile Protocol," https://core.telegram.org/mtproto, 2018.

[5] A. D. Luca, S. Das, M. Ortlieb, I. Ion, and B. Laurie, "Expert and Non-Expert Attitudes towards (Secure) Instant Messaging," in *SOUPS*, 2016.

[6] R. Abu-Salma *et al.*, "Obstacles to the Adoption of Secure Communication Tools," in *IEEE S&P*, 2017.

[7] E. Vaziripour *et al.*, "Is that you, Alice? A Usability Study of the Authentication Ceremony of Secure Messaging Applications," in *SOUPS*, 2017.

[8] J. Wu and D. Zappala, "When is a Tree Really a Truck? Exploring Mental Models of Encryption," in *SOUPS*, 2018.

[9] J. Tan *et al.*, "Can Unicorns Help Users Compare Crypto Key Fingerprints?" in *CHI*, 2017.

[10] S. Dechand, A. Naiakshina, A. Danilova, and M. Smith, "In Encryption We Don't Trust: The Effect of End-to-End Encryption to the Masses on User Perception," in *IEEE EuroS&P*, 2019.

[11] A. diSessa, *Models of Computation: User CenteredSystem Design: New Perspectives on Human ComputerInteraction*. Lawrence Erlbaum, 1986.

[12] J. Wu *et al.*, "'Something isn't secure, but I'm not sure how that translates into a problem': Promoting Autonomy by Designing for Understanding in Signal," in *SOUPS*, 2019.

[13] E. F. Foundation, "Communicating with Others," in *Surveillance Self-Defense*, 2018.

[14] Citizen Lab, "Analysis of End-to-End Encryption in LINE," in *App Privacy and Controls*, 2017.

[15] D. J. Solove, ""'I've Got Nothing to Hide" and Other Misunderstandings of Privacy," *San Diego Law Review*, vol. 44, pp. 745–772, 2007.

[16] S. Gaw, E. W. Felten, and P. Fernandez-Kelly, "Secrecy, flagging, and paranoia: Adoption criteria in encrypted email," in *CHI*, 2006.

[17] K. Krombholz, K. Busse, K. Pfeffer, M. Smith, and E. Zezschwitz, "'If HTTPS were secure, I wouldnt need 2FA': End user and administrator mental models of HTTPS," in *IEEE S&P*, 2019.

[18] W. Tong, S. Gold, S. Gichohi, M. Roman, and J. Frankle, "Why King George III Can Encrypt," http://randomwalker.info/teaching/spring-2014-privacy-technologies/king-george-iii-encrypt.pdf, 2014.

[19] A. Demjaha, J. Spring, I. Becker, S. Parkin, and A. Sasse, "Metaphors considered harmful? An exploratory study of the effectiveness of functional metaphors for end-to-end encryption," in *USEC*, 2018.

[20] W. Bai, D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek, "An Inconvenient Trust: User Attitudes toward Security and Usability Tradeoffs for Key-Directory Encryption Systems," in *SOUPS*, 2016.

[21] W. Bai, D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek, "Balancing Security and Usability in Encrypted Email," *IEEE Internet Computing*, vol. 21, no. 3, pp. 30–38, May 2017.

[22] N. Gerber, V. Zimmermann, B. Henhapl, S. Emeröz, and M. Volkamer, "Finally Johnny Can Encrypt: But Does This Make Him Feel More Secure?" in *ARES*, 2018.

[23] L. J. Camp, "Mental Models of Privacy and Security," *IEEE Technology and Society Magazine*, vol. 28, no. 3, 2009.

[24] F. Asgharpour, D. Liu, and L. J. Camp, "Mental Models of Security Risks," in *FC'07/USEC'07*, 2007.

[25] J. R. C. Nurse, S. Creese, M. Goldsmith, and K. Lamberts, "Trustworthy and Effective Communication of Cybersecurity Risks: A Review," in *STAST*, 2011.

[26] R. E. Mayer and R. B. Anderson, "The Instructive Animation: Helping Students Build Connections Between Words and Pictures in Multimedia Learning," *Journal of Educational Psychology*, vol. 84, no. 4, pp. 444–452, 1992.

[27] R. E. Mayer, *Multimedia Learning*, 2nd ed. New York, NY, USA: Cambridge University Press, 2009.

[28] N. R. Council, *How People Learn: Bridging Research and Practice*, M. S. Donovan, J. D. Bransford, and J. W. Pellegrino, Eds. The National Academies Press, 1999. [Online]. Available: https://www.nap.edu/catalog/9457/how-people-learn-bridging-research-and-practice

[29] J. R. Anderson, *Rules of the Mind*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc., 1993.

[30] "LINE," https://line.me/en/, 2018.

[31] Rakuten Viber, "Viber," https://www.viber.com/, 2018.

[32] "Threema," https://threema.ch/en, 2018.

[33] "ProtonMail," https://protonmail.com/, 2018.

[34] A. L. Strauss and J. M. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA, USA: Sage Publications, Inc, 1998.

[35] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology (second edition)*. Sage Publications, 2004.

[36] S. Barge and H. Gehlbach, "Using the Theory of Satisficing to Evaluate the Quality of Survey Data," *Research in Higher Education*, vol. 53, no. 2, pp. 182–200, 2012. [Online]. Available: http://www.jstor.org/stable/41349004

[37] R. Rosenthal, R. Rosnow, and A. Kazdin, *Artifacts in Behavioral Research: Robert Rosenthal and Ralph L. Rosnow's Classic Books*. Oxford University Press, 2009.

Appendix A.
Quiz Questions

Suppose TextLight is a messaging app which has end-to-end encryption (E2EE). You send a message to your friend, Bob, using TextLight, to invite him to watch the movie *Star Wars* tonight. Please answer the following questions based on your understanding of end-to-end encryption (E2EE). [Note: all questions except Q9 were five-point Likert-scale choice questions: "Very difficult," "Somewhat difficult," "Neutral," "Somewhat easy," "Very easy."]

1) Suppose Dave could observe the communication between your phone and the TextLight company, how difficult would it be for Dave to learn the contents of your invitation message?
2) How difficult would it be for the TextLight company to learn the contents of your invitation message?
3) Consider your Internet or mobile service provider, such as Comcast, Verizon, AT&T, Sprint, etc. How difficult would it be for these service providers to learn the contents of your invitation message?
4) If a hacker, Chuck successfully steals data that is stored in the computers at the TextLight company, how difficult would it be for Chuck to learn the contents of your invitation message?
5) Some government intelligence or national security agencies (e.g. NSA) request the TextLight company to hand over stored user messages to them. How difficult would it be for these agencies to learn the contents of your invitation message?
6) Suppose George has previously installed some malware on your phone which could record your activities, how difficult would it be for George to learn the contents of your invitation message?
7) How difficult would it be for someone (for example, a hacker) to modify messages between you and Bob during your conversation, so that Bob will receive an invitation to watch the movie *Jurassic Park*, instead of *Star Wars*?
8) Suppose both your and Bob's Textlight usernames and passwords are **not** hacked/stolen/guessed by other people. How difficult would it be for someone (for example, a hacker) to impersonate Bob in order to respond your invitation message? For example, you receive a response message that seems to be from Bob, but was actually sent by Frank.
9) Please sort the following tools based on how secure you feel using them, with the most secure tool at the top.
   a) Sending messages using SMS text messaging app
   b) Making a direct landline/mobile phone call
   c) Sending Email
   d) Sending messages in instant messaging apps with E2EE
   e) Sending messages in instant messaging apps without E2EE

Appendix B.
Interview Questions

Now we will ask you some questions about how you feel about these slides.

1) Could you briefly recall what you have learned from our tutorial?
2) How difficult or easy to understand was our tutorial? [Very difficult, difficult, neutral, Easy, Very easy]
3) How informative was our tutorial? [Very non-informative, non-informative, neutral, informative, very informative]
4) What piece of the tutorial you think is most important to you to help you learn E2EE? Why? What about non-important parts?
5) What part of the tutorial you think is most surprising to you? Least surprising? Why?
6) Suppose you want to teach E2EE to your family or friends who don't know E2EE before, what part of the tutorial do you want them to learn, or should they know? Why?
7) What part of the tutorial you like or dislike, think (not) clear, or (not) well-articulated?

Appendix C.
App Message Critique

Nowadays, some messaging apps also have E2EE. They also try to write some short introductions to their users. Here are some of the examples. Please take some time to read them. We anonymized these app names, and name them all as TextLight.

C.1. App Messages

APP 1 - Line: Letter Sealing is a feature that provides end-to-end encryption (E2EE) for chat room messages. E2EE is a communication system designed so that messages saved on our servers are encrypted and cannot be read by anyone except the sender and receiver of the message. Letter Sealing uses unique, user-specific encryption keys which allow users to safely and securely send messages to one another.

APP 2 - WhatsApp: Many messaging apps only encrypt messages between you and them, but TextLight's end-to-end encryption ensures only you and the person you're communicating with can read what is sent, and nobody in between, not even TextLight. This is because your messages are secured with a lock, and only the recipient and you have the special key needed to unlock and read them. For added protection, every message you send has its own unique lock and key. All of this happens automatically: no need to turn on settings or set up special secret chats to secure your messages.

APP 3 - Telegram: All messages use end-to-end encryption. This means only you and the recipient can read those messages – nobody else can decipher them, including us here at TextLight. TextLight can help when it comes to data transfer and secure communication. This means that all messages that you send and receive via TextLight cannot be deciphered when intercepted by your

ISP, network administrator or other third parties. But please remember that we cannot protect you from your own mother if she takes your unlocked phone without a passcode. Or from your IT-department if they access your computer at work. Or from any other people that get physical or root access to your phones or computers running TextLight.

APP 4 - Viber: On TextLight, end-to-end encryption is always turned on. Encryption keys exist on user devices and nowhere else. So no one but you and the people you're communicating with can see your messages or hear your calls, not even TextLight. The messages you send make their way from your device to the recipient in the form of a code that only the recipient's device can translate to plain text.

When your chats are protected by end-to-end encryption, no one has access to them. Perhaps most importantly, TextLight doesn't have access to them – which means nothing you share can be used to target you later. TextLight can't share what it doesn't have and, since TextLight doesn't have access to the content of your conversations, it can't share it with third. So, you can be sure you won't mysteriously start seeing ads related to something you were just talking about with a friend on TextLight.

APP 5 - Threema: TextLight end-to-end encrypts all your messages. All encryption and decryption happen directly on the device, and the user is in control over the key exchange. This guarantees that no third party – not even the server operators – can decrypt the content of the messages and calls. Only the intended recipient, can read your messages.

TextLight uses two different encryption layers to protect messages between the sender and the recipient.

- End-to-end encryption layer: this layer is between the sender and the recipient.
- Transport layer: each end-to-end encrypted message is encrypted again for transport between the client and the server, in order to protect the header information.

The crucial part is that the end-to-end encryption layer passes through the server uninterrupted; the server cannot remove the inner encryption layer.

APP 6 - ProtonMail: In TextLight, your data is encrypted in a way that makes it inaccessible to us. Data is encrypted on the client side using an encryption key that we do not have access to. This means we don't have the technical ability to decrypt your messages, and as a result, we are unable to hand your data over to third parties. With TextLight, privacy isn't just a promise, it is mathematically ensured.

When you use E2EE to send a message to someone, no one monitoring the network can see the content of your message – not hackers, not the government, and not even the company (e.g. TextLight) that facilitates your communication.

This differs from the encryption that most companies already use, which only protects the data in transit between your device and the company's servers. For example, when you send and receive a message using a service that does not provide E2EE, the company has the ability to access the content of your messages because they also hold the encryption keys. E2EE eliminates this possibility because the service provider does not actually possess the decryption key. Because of this, E2EE is much stronger than standard encryption.

C.2. Interview Questions

1) Generally speaking, what do you think of these introductions? [Followup: Do you think these introductions can help general users understand E2EE better?]
2) Please use the green pen to mark the parts you think the introductions are good, and use the red pen to mark the parts you think are not good. [Followup: Why do you think these parts are (not) good?]
3) What else you might like these introductions to tell you?
4) Please choose the best one from these introductions.

Appendix D. Message Design

Now you have learned our tutorial about E2EE, and have read some tutorials used by some messaging apps. Could you design a short introduction if you want to teach other people about end-to-end encryption? We don't expect you to come up with a long document describing every aspect of E2EE. Instead, please pick the most important pieces you want to convey. Try to come up with the introduction within 200 words [1].

D.1. Participants' Designed Messages

- **P1:** All messages use end-to-end encryption. This means only you and the recipient can read those messages - nobody else can decipher them, not even TextLight. All messages that you send and receive via TextLight cannot be deciphered when intercepted by your ISP, network administrator or any other third parties. Unless your usernames and passwords are stolen or guessed, there is no way that your messages could be modificated by a third person, nor the impersonation could happen. Nothing you share can be used to target you later. But please note that end to end encryption cannot protect your messages from any other people that get physical or root access to your phones or computers running TextLight.
- **P2:** E2EE is exceptionally important in protecting your information. Your information is exposed every time you utilize your mobile device. There are three potential exposures, the service providers, the app company, and hackers. In efforts to protect yourself against these parties please ensure your devices has E2EE protection. This protection allows you information to remain secure not even the company itself has an encryption key to decode, grant access to other and leave your personal information unsecured. Additionally, one must take precautions against those who shoulder surf and readily giving information/ access to others (giving passwords or forwarding information). E2EE is the best method of protection and all who utilize text messaging services should utilize its services.

1. We copied the exact words written by the participants, including their typos.

- **P3:** End-to-end encryption (E2EE) is a process that protects your messages from third parties including but not limited to an individual, your app company or internet service provider. As a result, only the sender and intended recipient of a message can view its contents. E2EE also protects messages from being duplicated, impersonated or modified by third parties. Although E2EE is a step forward in protecting user information, app companies are still researching ways to secure user information when Malware software is detected.
- **P4:** E2EE is an encryption used for securing your messaging. It sends an signal through your internet, sends to whom your messaging. With a decrypted message only they can see and open.
- **P5:** With other, non-encrypted messaging apps, your messages might be vulnerable to hacking before reaching your recipient. With TextLight, your message is automatically encrypted with a unique lock and key, ensuring only you and your recipient can read what is sent. We use two layers of encryption so that no one else– not even server operators or TextLight employees– can access your message.
- **P6:** Do you ever wonder who may get access to your private messages/texts? Nowadays, it is very easy for hackers to get your private information, but NOT if you use TexLight app! TexLight is a specially-encrypted app that ensures that only you and your intended recipient can view the messages. This encryption is so secure that even the TexLight company cannot decrypt the messages, so you can feel secure in your communication. Try out yourself by downloading the app here.
- **P7:** Many apps, text messages, and even phone calls are not secure these days. The government, hackers, and even phone companies can have access to the content of your messages. People can even modify messages that you send in between you and your friend. Thus, it is essential to find ways to provide more security when communicating over the web. One way people do this is through E2EE. It is a secure messaging app that makes it almost IMPOSSIBLE for people to gain access to the content of your message. For example, a messaging app that is E2EE may know that you have sent a message but they do not know what the content is. This also eliminates risk of hackers. Even if a hacker were to hack the company, they would not be able to alter or gain access to your content. It is actually more secure than other platforms such as text messages, imessage, emails, or even phone calls. It eliminates a lot of risk. However, there are still some risks that can not be eliminated such as someone looking over your shoulder while you're texting. However, overall, it is a more secure way for people to communicate via the web.
- **P8:** E2EE is a system which will allow for the safe and secure transmission of messages. This will make it such that the message can only be viewed by the sender and the intended recipient.
- **P9:** The advantage of using end-to-end encryption (E2EE) is that it prevents your messages from being read when they are in transit and when they are stored by the app company. Non-E2EE approaches can prevent messages in transit from being compromised, but they are still at the risk of being exposed to third parties if the app company itself compromised. For example, non-E2EE cannot prevent the message from being exposed to:
  1) Rogue employees
  2) Requests by government agencies
  3) Hackers who find a way to access the app company's servers

  However, the downside is that both E2EE and non-E2EE apps are still at risk when it comes to the user's phone. Such cases include when malware is installed or another person looks over the user's shoulder as they type in their passwords.
- **P10:** Hey everyone, I want to share some important news about the safe and secure software that can protect you, and your information you pass to another. This software is called E2EE it provides end to end encryption protection with messages and calls are exchanged and not even the providers can see it. However if anyone is looking over your shoulder that the only possible way your information can be shared. Overall this software is definitely something to look into if you have privacy difficulties.
- **P11:** E2EE is a very useful and secure way of assuring its user's that the contents of their messages cannot be viewed or manipulated in any way, throughout the transit and retrieval of the message. Unlike messaging apps that choose not to use end-to-end encryption, messages with E2EE cannot even be accessed by the app company themselves. As amazing as it may seem, the only possible way for a hacker to access the encrypted information would be if they had the ability to install malware on one's phone.
- **P12:** E2EE is a user friendly feature which protects both the sender and receive, as a result of method the data is encrypted stored. The positive capabilities eliminates the possibilities of a hacker intercepting messages or restrict the company's service provider's from executing decryption keys to decipher content. With TextLight, the software makes it vertically impossible for unauthorized outsider to gain access since there's no monitoring of the network by the company or third parties.
- **P13:** End to end encryption is the most secure platform to convey your messages to the other user.It does not even let the network providers and the company itself to access to our messages. We have a few misconceptions about E2EE like we think that our information can be hacked or used by other agencies be it government or non-government but it cannot. Along with this, we think that mailing is the best and most secure method for communication but what I have found is that apps with E2EE are more secure that E-mails, phone messengers and phone calls. We just need to take care about the malwares installed in our phone and if any other person can have access to our phone.
- **P14:** The end-to-end encryption(E2EE) means only you and your intended recipients can read the content

of messages. Under non-E2EE situation, there are risks in which your message can be subject to deciphering, interception and modification. The internet service providers can get access to the content of your messages during transit, the app company can can read your messages and hand them to government agency when required. Moreover, a hacker can easily intercept and modify your messages. Using E2EE can eliminate the possibility of leaking the contents of your messages in situations mentioned before by adding a shield to your message so that only you and the recipient can open that shield. More importantly, no message modification and impersonation will take place. However, E2EE can't protect you from shoulder surfing, an unreliable recipient or malware installed on your phone.

- **P15:** End-to-end encryption otherwise as E2EE is a way to protect messages, phone calls, data transfers, and other similar things from being obtained. E2EE does this by essentially putting a shell that is impenetrable around the information being sent. This means that it is nearly impossible for someone to break the shell and get the information that the shell is protecting. The shell protects against everything such as the text messaging company, hackers who try to intercept the message, or internet service providers. The only time E2EE does not work is when a virus has already been placed on your phone/computer, someone gets into your phone/computer, or whoever you are sending the information shares it.
- **P16:** I learned that E2EE is an end-to-end encryption that secure messages between you and others. E2EE secures each messages with a key that you have with each message sending out. Non E2EE is more vulnerable to hackers that can intercept, edit, & view your outgoing messages. Government officials can legally obtain messages from the app company or service provider no matter how secure the encryption is. Sometimes message can also be lost in transit.
- **P17:** When you use E2EE to send a message to someone, E2EE encrypts the data in a way that makes it inaccessible to third parties, as the message is being delivered. This means that the app company, internet service providers (such as comcast or verizon), government agencies, and even hackers do not have the ability to decrypt your messages. Additionally, the app is unable to hand your data over to third parties. No one monitoring the network can see the contents of your message.
- **P18:** End-to-end encryption, or E2EE, protects messages in that it only allows you (the user) and your intended audience to read the message. Data is protected on the user side using an encryption that the messaging app company - or any other third party (including hacker, the government, impersonators, etc.) - have access to, therefore no one else can see the content of the message. E2EE does not, however, protect receivers if malware has been installed in their phone.
- **P19:** There are more benefits than risks to using E2EE, or End-to-End Encryption. E2EE mitigates some of the risks message senders are exposed to in every day life while texting their friends, sending invitations through Whatsapp, or communicating in other ways via instant messaging.
  By using E2EE, this process creates a lock and key for your messages. When you send a message using E2EE, the message travels from your Internet Provider, through the messaging app servers, through the recipient's Internet Provider, and to the recipient. The lock prevents hackers, rogue employees, and the app company/IP company themselves from viewing the messages, and only the recipient has the key to unlock these messages. By using E2EE, you decrease the likelihood that someone can hack into your phone. While there are benefits to using E2EE, every day risks still pose a threat, like a thief stealing your phone, someone looking over your shoulder, or if someone installs malware on an unlocked phone.
- **P20:** End-to-end encryption is an encryption technique that utilizes user-specific encryption keys that only the sender and receiver of messages have access to. This encryption technique ensures that messages from one party to another are secure during transit between parties and when the company that stores the data for these messages is compromised. The algorithms that are utilized here are extremely strong and would take third parties an impossible amount of time to crack and decrypt. End-to-end encryption is an essential tool to protect users' privacy in the modern age, whether it be from the government or private hackers.
- **P21:** E2EE enables a more secure form of communication, making users less vulnerable to security threats. Therefore, when a user is incentivized to have secure communication, apps using E2EE should be utilized over non-E2EE apps. While this is a general principle, it's worth noting that messages sent over E2EE are still vulnerable in some regards (just fewer parts than when not using E2EE).
- **P22:** In TextLight, your data is protected with end-to-end encryption in a way that no one except you and the person you are communicated with can read your messages. This is because each message you send has its own encrypted lock and key so that not only us (the company) but also the internet service providers and even the government CANNOT decrypt your messages and hand over to third parties. Also, the encryption works automatically and hence there is no need to turn on settings or set up special secrets chats.
- **P23:** End-to-end encryption means that each user has a key which only their phone has. When travelling between service providers and the app company, the message is encrypted and cannot be read until the key unlocks the message. With normal message services, they hold the unencrypted messages which could be read at any time, and third parties such as hackers or the government could see the contents. The encryption algorithms are also very strong (so that no hacker can break it), so the only way for somebody besides the intended recipient to view the message contents is if their username or password is stolen/revealed.
- **P24:** E2EE is a tech that can protect your messages from any other third party by controlling E2EE pro-

cess happening directly on the device, so that only the intended recipient can read your messages.

- **P25:** E2EE is a great way to ensure that your messages will be securely transmitted to the recipient of the message originally intended for. The data is encrypted where no one has directed access to the text contents but the sender and the receiver of the information. Not even the monitoring company, hackers and the government can see the message being sent in the secure way. On particular E2EE apps the encryption is always turned on the be coded and will decode once the message has been fully received by the recipient. This E2EE protects against message modification and impersonation. Not even usernames and or passwords can be stolen or guessed.