“If You Want to Encrypt It Really, Really Hardcore…”: User Perceptions of Key Transparency in WhatsApp

Konstantin Fischer
Ruhr University Bochum
konstantin.fischer@rub.de

Markus Keil
Ruhr University Bochum
markus.keil@rub.de

Annalina Buckmann
Ruhr University Bochum
annalina.buckmann@rub.de

M. Angela Sasse
Ruhr University Bochum
martina.sasse@rub.de

Abstract

WhatsApp is the first popular chat app to roll out a real-world, large-scale implementation of key transparency. If implemented correctly, key transparency allows users to check whether they are currently victim of a Machine-in-the-Middle attack mounted by WhatsApp server operators. Through 16 in-depth semi-structured interviews with WhatsApp users in Germany, we investigate how people judge and perceive the security and privacy of chat apps, whether end-users perceive benefits from key transparency, and how this affects trust and usage.

We find that our interview participants mostly know what end-to-end encryption is, but that they struggle to show an understanding of the nuanced threat models needed to grasp the point of key transparency. Seeing key transparency in action led to a slight increase in perceived security in some, while others dismissed it as an unconvincing UI sham that would not change their presumptions about WhatsApp and its companies’ motives. Some participants even felt less secure after performing a key transparency check, which we attribute to certain misconceptions we uncovered during the interviews.

We conclude that exposing end-users to key transparency, without an accompanying explanation, is unlikely to directly meaningfully enhance trust or perceived security, and can even lead to users feeling less secure in some cases. We underline that the real strength of KT lies in 1) what we call the "deterrence effect" and 2) the future possibility to better automate key transparency checks. Based on our results we offer recommendations for industry practitioners as well as for promising future work in academia.

Keywords

Key Transparency, Chat Apps, Privacy, Usability and HCI

1 Introduction

WhatsApp has over 2 billion users and is the most used chat app in the world [39]. WhatsApp implemented end-to-end encryption (E2EE) in 2016, which could be argued was one of the greatest privacy improvements for personal messaging “over night”. However, WhatsApp uses only opportunistic end-to-end encryption [23] by default, which allows for server-side Machine-in-the-Middle (MitM) attacks. For authenticated end-to-end encryption, i.e., a connection where MitM attacks are ruled out, extra steps are necessary. Most popular E2EE chat apps offer authentication ceremonies, which allow users to check that there is no MitM attack present [23]. These authentication ceremonies have been shown to be unknown or unusable for most users in many previous works[3, 20, 23, 24, 38]. Researchers proposed an automated alternative to detect if service operators behave dishonestly: Key transparency (KT)—a promising security feature that would allow detecting server-side MitM attacks without the need for user effort [10, 29, 34]. WhatsApp are the first popular chat app to roll out a real-world, large-scale implementation of key transparency [27]. In its current implementation in WhatsApp, the “automatic” key transparency check has to be started manually by the user. In the app, the key transparency mechanism is surfaced to users[1] as some extra UI elements, as seen in Figure 1.

Little is known about how end-users perceive, understand, and interact with this new security mechanism, which, in its current implementation does require user interaction. Prior research has shown that users often misunderstand other encryption features or overlook them entirely. We lack insight into whether and how key transparency can shape trust in chat apps and security perceptions. Additionally, the last major work that focused on how users perceive the security of WhatsApp was published in 2019 (over 5 years ago) [17], users’ perceptions might have changed. This paper explores how everyday WhatsApp users assess the app’s security, whether they understand or value key transparency, and what misconceptions may influence their perceptions, with the aim of ultimately informing design decisions in future iterations of key transparency implementations.

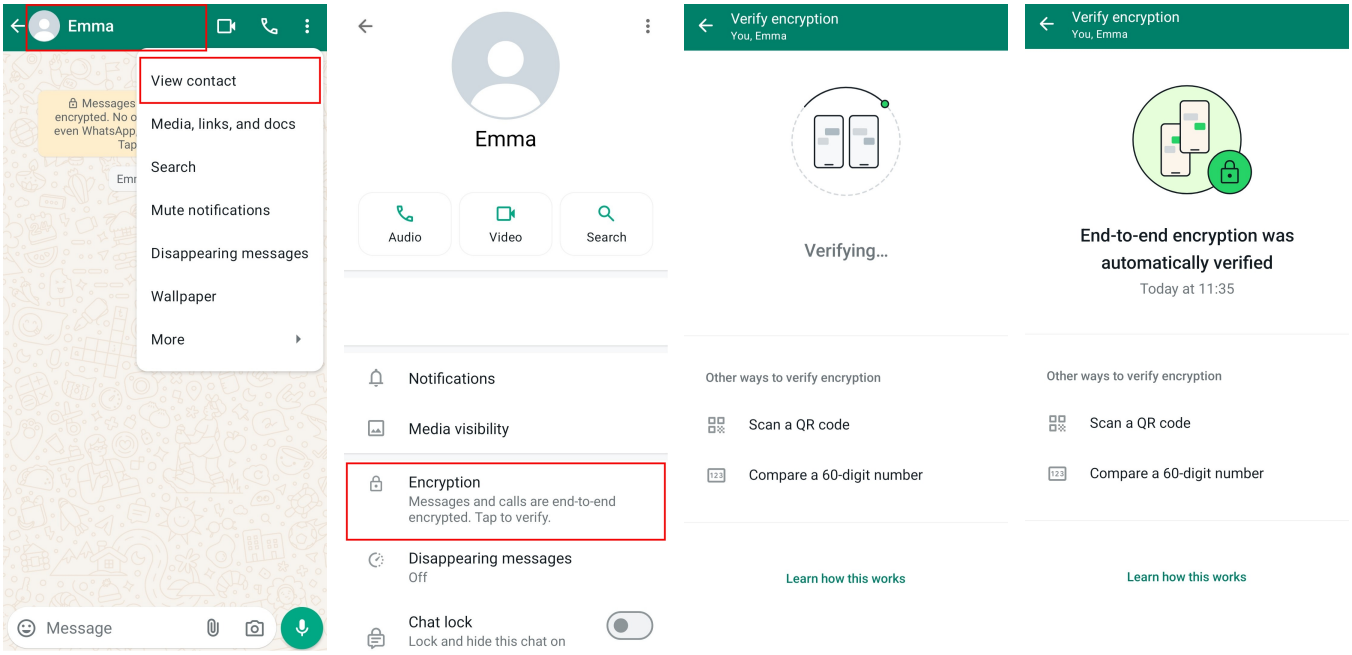
To do this, we first provide an overview of the current threats to chat app end user message privacy, then highlight where key transparency fits in as a response to one of these threats, and then, on the basis of 16 semi-structured interviews with WhatsApp users, explore the following research questions:

RQ1 How do end-users perceive and assess the security of WhatsApp?
RQ2 Do end-users perceive any benefits from key transparency in WhatsApp?
RQ3 What are misconceptions about encryption and key transparency in WhatsApp?

[1] At the time of writing, the key transparency UI is only available in WhatsApp Messenger for Android (February 2025)

Proceedings on Privacy Enhancing Technologies 2025(4), 1039–1054

https://doi.org/10.56553/popets-2025-0170



(a) (b) (c) (d)

Figure 1: The Key Transparency User Interface in WhatsApp. (a) Chat screen with highlighted options to switch to the contact screen. (b) Contact screen with highlighted option to switch to the "Verify Encryption" screen and immediately start KT lookup. (c) "Verify Encryption" screen with pending KT lookup. (d) "Verify Encryption" screen with completed KT lookup.

Ethics. Our institution has no formal IRB process for computer science studies. Both studies were designed with strict ethical considerations according to the Menlo report [25] and the GDPR. Participants were compensated monetarily for their efforts. The collected data was anonymized, and participants were provided with a consent form that had all the information on how and why data is recorded and stored. The participants were informed about withdrawing their data during or after the study.

2 Background

We give a short overview of security features and threats when using WhatsApp and define the terms security, privacy, and trust.

2.1 Threat Model for Message Confidentiality

In April 2016, WhatsApp deployed *end-to-end encryption* (E2EE) to their service, for all messages, by default [30]. In a technical white paper WhatsApp defines E2EE as: "*communications that remain encrypted from a device controlled by the sender to one controlled by the recipient, where no third parties, not even WhatsApp or our parent company Meta, can access the content in between.*" [48]. WhatsApp claims to employ the *Signal* protocol [30], which uses standard modern cryptographic primitives and the Double-Ratchet Algorithm to provide confidentiality, integrity, forward secrecy, future secrecy, as well as other advanced security properties [21, 37, 44], and has been proven secure in formal analyses in different ways [4, 7, 12, 13].

2.1.1 Chat App Threat Model. When using a modern chat app like WhatsApp, there are different ways in which message confidentiality can be breached.

Network Threats Classic network attacks like sniffing or routing-based MitM attacks on local networks are prevented by TLS encryption and authentication, including certificate pinning, for all connections made from user devices to the WhatsApp servers.

End-User Device Threats Chat messages are stored on the end-user devices. If a device is compromised by, e.g. malware, or if the device is physically taken and the device unlock PIN was acquired via shoulder surfing or forced disclosure by, e.g., law enforcement, message confidentiality is breached. Message confidentiality can also be breached if a physically close person simply peeks at the user's display.

Cloud Backup Threats Modern smartphone operating systems offer easy backup solutions of app data. These backups may then be accessible to the given cloud storage service provider, or protected by an extra layer of encryption like Apple's iCloud *Advanced Data Protection* [6, 43]. WhatsApp additionally allows creating chat backups from within the app and storing them on the platform's cloud storage providers (Google Drive or Apple iCloud). Since 2021, WhatsApp allows users to opt-in to "end-to-end encryption" of these backups [16].

Client-Side Code Threats Users install and run the closed-source WhatsApp application on their phones. If the app has intentional backdoors or unintentional security vulnerabilities, message confidentiality can be breached.

Honest-But-Curious Operator Most chat apps rely on central servers to relay messages between end-user devices and distribute public keys. The default opportunistic end-to-end encryption in WhatsApp protects message contents from honest-but-curious [36] operators.

Malicious or Compromised Operator Opportunistic end-to-end encryption means that users rely on the server operator to distribute the correct public keys to end-user devices. An active server-side attacker could simply distribute attacker-controlled keys instead, and thus mount MitM attacks arbitrarily, breaching message confidentiality. While this can not be prevented, these MitM attacks can be detected by users performing authentication ceremonies or key transparency checks. Note that there is no public evidence that WhatsApp has ever tried to mount a MitM attack on their users.

We underline that key transparency, if implemented correctly, is only capable of protecting against the threat of a malicious or compromised operator. End-user device threats, cloud backup threats, and client-side code threats persist. We discuss this further in subsubsection 2.2.1.

2.2 Key Transparency in WhatsApp

WhatsApp's implementation of key transparency is mainly based on the proposed key transparency protocol Parakeet [29], which is in turn based on SEEMless [10] and CONIKS [34].

To enable key transparency, the WhatsApp server publishes a cryptographic commitment for an auditable key directory that links identities (phone numbers) to public keys. User devices can then ask the server for inclusion proofs to check the public keys they received against the global commitment, i.e. check that WhatsApp serves the same key for the same identity to everyone. The emerging security property is called *key consistency* and prevents equivocation attacks, i.e. a server presenting different key directory states to different users. To handle key directory updates, at every epoch (currently every minute), WhatsApp publishes a new commitment, including all changes to the key directory since the last epoch and the last commitment. Third party witness *CloudFlare* then verifies 1) the *append-onlyness*, i.e. that the key directory has only undergone allowed changes like adding keys for new identities, or adding updated keys to existing identities, but never removing recent key updates from the directory, using an inclusion proof that is published with each commitment and 2) the *uniqueness*, i.e. that there is only one commitment for the current epoch. If both checks succeed, CloudFlare cross-signs the epoch as a third party witness[29, 35].

Starting a KT check in WhatsApp does two things:

(1) Chat partner's key lookup: Checks that WhatsApp serves the same key to you for this contact as to everyone else. This would detect equivocation attacks.
(2) Personal key lookup: Checks that everyone who does the KT check receives the correct key for you (in that moment).

From this follows that, if both chat partners do a successful key transparency check, they can be certain that there is no MitM attack happening on their chat at that point in time.

Users can start a key transparency check in the WhatsApp user interface as depicted in Figure 1.

Parakeet and SEEMless propose options for clients to monitor their own key history, which would detect a MitM attack, even if the key were changed back to its original value in time for lookup requests to verify successfully again. At the moment, however, WhatsApp's implementation does not offer a key history lookup. Each lookup request not only checks the recipient's key but also the sender's own key as well. As long as Bob does not start a lookup request for any recipient, his client would not detect the attack. Additionally, key transparency, like authentication ceremonies, is a reactive system. Neither actively prevent attacks, but they can detect ongoing attacks.

2.2.1 Key Transparency Threat Model Implications. We want to put the security benefit gained by KT into the bigger context of the general threat model for chat app message privacy (cf. subsubsection 2.1.1). Key transparency, if implemented correctly, is only capable of protecting against the threat of a malicious or compromised operator. This is partly achieved by introducing "witnesses", i.e. a third party service as KT auditor and KT signing service. The result is an overall small, nuanced shift in the theoretical threat model. We try to visualize this nuanced shift with Table 1: After introducing KT to WhatsApp to ensure key consistency, users still have to trust the client-side code created by WhatsApp, and they now have to trust that the KT service is run correctly.

Partial trust in KT auditor and KT signing service. In KT as set up by WhatsApp, the KT auditor cryptographically verifies the append-only proofs, i.e. verifies that no key directory entries have been removed or rolled back. Since these proofs are public, there could be an arbitrary number of AKD auditors, and the end-user needs to trust that only one of them does not conspire with the service operator. The user could also verify the cryptographic proof themself, but this is computationally expensive and needs to be done for every epoch, i.e. every minute.

The KT signing service guarantees that there is only one valid global commitment for every epoch, by cross-signing not more than one commitment per epoch. This service's operators are chosen by WhatsApp, as they decide which cross-signatures the client-side code will accept as valid. If they conspire, the user themself could theoretically still validate whether WhatsApp only publishes one valid commitment per epoch to a public write-only AWS bucket [47], but this can not be expected from everyday WhatsApp users.

Server Operator Trust. We want to underline the subtleness of the difference in threat models for opportunistic and authenticated E2EE in chat apps, from an end-user's point of view: In a non-E2EE chat app like Telegram, Alice has to trust the service operator not to read her messages. In an opportunistic E2EE chat app like WhatsApp, Alice has to trust the service operator to distribute the keys honestly, so that they can't read her messages. If Alice does not trust WhatsApp to distribute the keys honestly, Alice can start a key transparency check—for which Alice has to trust that

Table 1: Chat app system components that need to be trusted by users to assume message security.

Security Architecture	*Client-Side Code*	*Message Relay Server*	*Key Distribution Server*	*KT Auditor*	*KT Signing Service*
1. Transport Encryption	●	●	N/A	N/A	N/A
2. Opportunistic E2EE	●	○	●	N/A	N/A
3. E2EE + Key Transparency*	●	○	○	◐	◐
4. E2EE + Auth. Ceremonies	●	○	○	○	○

● - trust required, ◐ - partial trust required, ○ - no trust required
*Note that in WhatsApp's implementation, the KT check has to be started manually by both chat partners, and there is only one 3rd-party AKD auditor and signer.

WhatsApp has implemented the key transparency service honestly and as advertised, on both server-side and client-side.

A less nuanced end-user point of view could lead to the conclusion that in any case, i.e. for any security Architecture described above, they simply *have to trust WhatsApp* to assume message confidentiality.

We additionally stress that, because WhatsApp is closed-source, users who care about message confidentiality have practically no option but to trust WhatsApp to implement their privacy features (E2EE, authentication ceremonies, Key Transparency) as advertised, and free of backdoors on the client side.

2.3 Foundational Definitions

Security. For this work, we define security as the extent to which data confidentiality and data integrity are protected against unauthorized access.

Privacy. We define privacy as the extent to which data collection, data sharing with 3rd parties, and data misuse are minimized.

Trust. We use the definition for trust given by Mayer et al. [32], who state that trust is "*the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party.*". Mayer et al. state that 1) the trustor's *propensity to trust*, and 2) the trustor's assessment or estimation of the trustee's *ability*, *benevolence*, and *integrity* leads to trust.

3 Related Work

We summarize key works in usable security research on end-to-end encryption, authentication ceremonies, and trust in chat apps, as well as prior research on the topic of key transparency.

3.1 Authentication Ceremonies

To detect MitM attacks on a chat, most chat apps allow users to perform authentication ceremonies. Previous work on authentication ceremonies identified a plethora of usability issues, including lack of feedback and text-heavy documentation in OTR for Pidgin [40], unintuitive feedback and inconsistent interfaces in ChatSecure [5], missing or unhelpful key reset notifications, misleading encryption status indicators, and poor guidance on the authentication process itself across WhatsApp, Signal, Viber, and Telegram [23, 38].

Studies that tried to teach users that authentication is necessary via intervention texts [46] or a more guiding UI [45] saw an increased number of performed authentication ceremonies in their test runs, but the participants still displayed poor understanding of the security implications and the need for authentication ceremonies in general. In a 2020 meta-analysis, Herzberg et al. [24] recommend re-framing the ceremony as a privacy check for high-risk users only, preventing them from exchanging messages until the ceremony is performed.In an 2023 extensive systematization of knowledge, Alatawi et al. [3] contrived and investigated a list of messaging apps, their E2EE features and authentication ceremonies. They conclude that all investigated apps are ineffective in repelling MitM attacks and that no app provided an effective and usable authentication ceremony.

Wu et al. [49] argue that the low risk of MitM attacks and the high response cost of authentication ceremonies make it a reasonable decision for everyday end-users to forgo manual authentication, even when threats and authentication processes are well understood. Fassl et al. [20] confirmed the high response costs with an auto-ethnographical study in 2023. They report cognitive load, forgetfulness, and social awkwardness as some of the biggest hurdles.

Notably, a number of the related works cited above mention key transparency as a potential solution to current usability troubles of authentication ceremonies.

3.2 Perceived Security of Chat Apps

In 2017, Abu-Salma et al. [1] conducted 60 interviews with end-users on E2EE in chat apps. End-users reported to believe a direct messaging app is secure if it has a large user base and high quality of service, and a majority of participants did not understand E2EE. The authors recommend leveraging secure tools with proved effectiveness instead of creating new ones.

Interviews with German WhatsApp users showed that interviewees' models of encryption were plagued with misconceptions and many believed that encrypted messages could still be eavesdropped by hackers, criminals, or governmental institutions [22] and that no technical solution could stop skilled attackers from getting their data [17]. Participants questioned WhatsApp's intentions behind the cost-free implementation of E2EE and articulated mistrust towards WhatsApp [22] and that, although WhatsApp informed its users of the introduction of end-to-end encryption prominently in the app in 2016, most were not aware of it [17]. We note that the authors of these works label "WhatsApp can read my messages" as a misconception—an assessment highly dependent on one's threat model (cf. subsubsection 2.1.1).

Studies on the effect of visibility of encryption showed that the wordings "encrypt" and "secure" perform comparatively well at increasing perceived security [18, 19], while in one study participants felt overall more secure not knowing about the objective of encryption at the same time [19]. Visual representations of encryption in

interfaces have been shown to have no or severely limited effects on perceived security [18, 41], understanding [18], and trust [41]. The authors recommend to increase trust in the chat app operator instead of studying visualizations of encryption further.

3.3 Key Transparency

The first formalization of an end-user key verification service for chat apps, i.e. *key transparency*, was CONIKS [34], built borrowing concepts of certificate transparency [26]. It consists of an authenticated data structure containing all users' public keys, which is periodically signed and published. By each client monitoring their own key entries in each published epoch, there is no need for any trusted third-party auditors. The idea is that the key directory ensures consistency while client-side monitoring guarantees correctness.

SEEMless [10], extends upon CONIKS by formalizing the notion of a *Verifiable Key Directory* (VKD), introducing the term *append-only zero knowledge set* (aZKS) and replacing the original data structure with the newly conceived *Persistent Patricia Trie* (PPTr), enhancing scalability by a considerable margin. Additionally, instead of having each client monitor their key revisions on each published epoch, users can monitor their key revisions at an arbitrary time, reducing monitoring costs for clients. This, however, introduces the need for a trusted third party auditor.

Parakeet [29] further tackles scalability problems of SEEMless with a direct vision to be using it for billions of users. An operation of reducing the data stored on the server by purging ancient and obsolete entries is introduced under the name "compaction". This loosens the requirements for append-only data structures and tackles the problem of storage requirements for data structures of this kind. Finally, Parakeet introduces a *consensusless* consistency protocol challenging the notion of *consensus* that the two previous works relied upon.

In 2023, Len et al. [28] present *OPTIKS*, a key transparency system heavily focused on scalability and performance. The authors also present a full server architecture and newly introduce *account decommissioning* and support for multiple user devices. Although OPTIKS does not fully match Parakeet in privacy, it claims better scalability while achieving the same level of security.

4 Method

We describe the methodology of our study, including a rationale for our study design decisions, participant recruitment, and the qualitative data analysis approach we used to address our research questions.

4.1 Study Design and Interview Guide

The main goal of our interview study was to better inform future KT UI design decisions by gaining insights into users' unbiased thoughts on the security of WhatsApp (RQ1) and their perceptions about key transparency in WhatsApp (RQ2+3).

Instrument Design Choices. We decided on a semi-structured interview format which allows for open-ended questions, enabling less biased responses, flexibility to explore opinions and ideas beyond our predefined scope and assumptions, space for participants to venture into topics not covered in the interview guide, and the freedom to follow up on interesting statements. For better ecological validity of our interview study results, we did intentionally *not* explain KT to our participants before having them use and judge KT in a scenario-based task, during which participants performed a key transparency check on their own phones with a newly added contact "Emma". We aimed to catch their impressions as they would be formed when they encounter the KT UI in the real world, by using WhatsApp. To induce a security-task mindset in the participants, we lent a strategy from previous work by Vaziripour et al. [45, 46] by having Emma ask the participants to transmit her credit card number. We then inquired about their feelings and understanding of the key transparency check they performed. The full interview guide is provided in Appendix B.

Interview Guide Structure. The interview guide included warm-up questions to facilitate rapport, followed by inquiries about WhatsApp's security, threat models, and the participant's trust in the WhatsApp company. We then introduced a task scenario in which the participant's "good friend Emma" just bought a new phone with a new phone number, asking the participant to add Emma as a contact on their phone and to start the *automatic encryption verification*[2], before sending sensitive credit card info to Emma. The interview then continues with questions about the participant's perceptions of the key transparency system, including perceived security, trust, and usability. Additionally, questions on the participant's privacy habits in WhatsApp and about WhatsApp's integrity are asked. The interviews ended with a short debriefing, where we answered all questions our participants had on chat app security. The full interview guide is provided in Appendix B.

Data Analysis. The interview recordings were transcribed using a "clean verbatim" style, and stripping any identifying information in the process. We used an iterative semi-open coding approach [9, 14, 42] to perform thematic analysis [11] for all interview transcripts. Two researchers coded the first six interviews independently, then compared and discussed their individual codebooks, merging and re-adjusting until they arrived at one common codebook. This codebook was then used to iteratively code the remaining 10 interviews. Our approach does not require the reporting of inter-coder agreement, as we resolved each conflict when it emerged, following established practices in the HCI community [33]. We provide the merged codebook in Appendix C.

4.2 Recruiting and Sample

Recruiting. The Interviewees were recruited using convenience sampling among the authors' friends, extended family, and friends of friends, as well as through university mailing lists and flyers posted to several university-related Facebook and WhatsApp groups groups.

A short online screening questionnaire (cf. Appendix A) allowed us to select participants reflecting a diversity in age, gender, occupation, and self-reported smartphone competence. We interviewed to saturation (P13-P16 added no new themes). Ultimately, 8 participants were recruited from authors' friends and family, and 8 via

[2]WhatsApp's UI does not use the term key transparency, but calls it "Automatic Encryption Verification" instead.

university mailing lists and flyers. We discuss the limitations of our recruiting method in subsection 4.3.

Participant Sample. The sample consisted of WhatsApp users aged 20–62 from diverse educational and professional backgrounds. The mean age was 42.5 years (Median: 27.5 years, SD: 17.2). They were 9 women, 6 men, and 1 non-binary person. Occupations covered a wide area, including students, accountants, teachers, nurses, and office administrators. All of them lived in Germany and used WhatsApp for Android in their daily lives. We invited only Android users to our interviews because WhatsApp on iOS does not offer the key transparency UI at this point in time. All participant quotes were translated from German to English.

At the time of our interviews, the current WhatsApp for Android version was *WhatsApp 2.24.16* (August 2024).

4.3 Limitations and Threats to Validity

Our use of convenience sampling introduces potential selection bias which might impact the external validity of our results. The reliance on personal and university-affiliated networks likely skewed our sample toward individuals with higher levels of education and digital literacy than the general population. This may help explain why many participants in our study expressed relatively vigilant or skeptical attitudes toward corporations and data protection practices, even when their understanding of underlying technologies was limited. Thus, our findings may not fully capture the perceptions and experiences of individuals with lower educational backgrounds or limited access to technical information.

While we aimed for diversity in age, gender, occupation, and smartphone competence, our sample (N=16) is not representative. All participants lived in Germany, which likely limits the generalizability of our findings across different cultural or regulatory contexts.

These limitations should be considered when interpreting our results, which are intended as exploratory insights rather than general claims. We encourage future research to replicate and expand on our findings using larger samples and including participants with lower digital literacy or non-German cultural backgrounds.

5 Results

To investigate and explore end-user perceptions of WhatsApp's security and key transparency in WhatsApp, we conducted semi-structured interviews with 16 WhatsApp for Android users.

5.1 Perceived Security and Trust in WhatsApp

In our interviews, we investigated whether users felt secure using WhatsApp and whether they trusted the WhatsApp company. We also inquired about their threat perceptions when using WhatsApp.

5.1.1 Perceived Security and End-to-End Encryption. After some warm up questions, we opened our interviews with a broad, exploratory open question: "Say... is WhatsApp secure?". The idea was to probe, as unbiased as possible, what end-users associate with security of chat apps in general. To our surprise, almost all of our interviewees, unprompted, mentioned WhatsApp's end-to-end encryption in their immediate response. P1 answered "Because you're always told there's end-to-end encryption, I think it's quite okay."

All our participants have heard the "end-to-end encryption" before, and most could explain the general concept correctly, i.e. that E2EE is supposed to prevent WhatsApp from reading chat messages.

Some interviewees showed confidence in WhatsApp's encryption practices

> *"The messages are definitely encrypted. And they can not be read by WhatsApp and not by anyone else. That's why I would say it is secure." —P14*

and deemed WhatsApp to be trustworthy: *"Maybe they *could* create something like a backdoor, but that sounds like a conspiracy theory to me."* (P6).

However, many were not convinced that WhatsApp's implementation of E2EE prevents WhatsApp, and sometimes even others, from breaking or circumventing the encryption: *"I believe that they have end-to-end encryption. But whether it's 100% secure, as they portray it—I don't believe that."* (P7). Even though our participants couldn't explain why, their gut feeling was spot on: WhatsApp implements E2EE, but this does not mean that there is absolutely no way WhatsApp could read their messages if they really wanted to (cf. subsubsection 2.1.1).

One participant stated even that *anyone* could gain access: *"WhatsApp says that it's end-to-end encrypted [...] but I think that at the end of the day, anyone can access it if they want it enough."* (P12).

Others were more sure:

> *"But every encryption can be reversed. [...] The way I imagine it is that if an encryption is created by some program or something, then a similar or different one would have to be able to decrypt that." —P3*

5.1.2 Trust in WhatsApp. When asked about their trust in the WhatsApp company, our participants often replied with how they would use WhatsApp, how secure they perceived the system to be, and some reported their specific threat models in regards to using WhatsApp. We learned that these concepts are highly intertwined in our participants' heads. We identified three main themes in partipants' responses:

Limited Trust in WhatsApp. One part of our interviewees argued that the fact that they use WhatsApp must mean that they trust it, but then were quick to qualify the extent of that trust: *"[My trust] does exist. I mean, I use WhatsApp."* (P3). One participant described their usage of WhatsApp to be despite a more general distrust: *"I don't think that I would trust Meta. But, if we're being honest, I trust them enough to use [WhatsApp]."* (P6). These participants said they trust WhatsApp *"enough to use it."* (P7), mostly meaning for everyday conversations.

High Trust in WhatsApp. Other interviewees reported to trust WhatsApp. They gave different reasons as to why or how far they would trust the company. First, one participant stated: *"Well in my day-to-day life I don't think about this. So I think [my trust in WhatsApp is] very high."* (P4). Other participants further specified how far they would trust the company: *"'I am trusting and trust them to not read my chats on purpose.' —P11"*, implying that their

chats might still be read by WhatsApp accidentally, or that they may let someone else gain access unwillingly or on purpose.

Some participants said they trust WhatsApp and justify this with its large user base: *"Because of the many users..."* (P8).

No Trust in WhatsApp. Most participants stated to not trust WhatsApp. It was commonly expressed that any claims made by WhatsApp or Meta regarding security could only be truly "trusted" if they were verified by trusted third parties in the form of experts, other companies, or their peers with more tech knowledge:

> *"I would actually wait for the feedback of an expert and then say 'Hey, is this really secure? Are you sure that I can send my bank details on there?'" —P15*

As reasons for low trust we found participants who had doubts about WhatsApp's *benevolence*: *"They could, if they wanted to, invest a lot of money in this to make it secure enough that no one else in the world could decrypt it. I doubt that they do that, though."* (P3). Reasons for the general lack of confidence in the WhatsApp company, among others, were past scandals: *"My trust is not very high, just because there was that Cambridge [Analytica] scandal and it didn't really win my trust."* (P5), and, most notably, a general bad reputation: *"But every time I hear anything, I've heard very critical things about Mark Zuckerberg and the Facebook corporation. Which belongs to Meta."* (P10). One participant added that, because of the network effect, they have no other feasible choice than to use WhatsApp:

> *"I think because Meta or WhatsApp is so big and everyone uses it that it doesn't necessarily make a difference for what I use. I think I would just as well trust Twitter if everyone in my social circle would be using that." —P1*

. We want to underline this sentiment: Multiple participants signaled that they effectively rely on WhatsApp, whether they trust it or not. P14 explained their view on WhatsApp's accountability and the idea that due to its many users, any misbehavior would be likely detected:

> *"It's a kind of institutional trust, that I trust that if something were wrong, the consumer protection agency's attention would be drawn to it. So not really the trust in the company but rather the trust that if they were to claim wrong things then that would be detected." —P14*

They feel secure sharing sensitive information via WhatsApp, even though they don't trust the WhatsApp company.

5.1.3 Threat Models. We asked our interviewees about potential security and privacy risks when sending messages on WhatsApp and outline the threat models we inferred from their responses. We report adversaries, assets, and attack vectors.

Attacks by WhatsApp or Meta. Most participants named WhatsApp or Meta as a threat actor. Some participants believed that WhatsApp could break their own end-to-end encryption *"I believe that there is a gap somewhere and [the encryption] apparently doesn't work"* (P2), others stated that the service provider would have ways to circumvent it *"Meta can probably... if they really wanted to they can probably read along."* (P6), and one assumed that all chat messages were stored by WhatsApp in a database: *"'The way I imagine it is that they have a database. Of all the things that are sent.' —P3"*. Other's mentioned meta data collection as a privacy risk. Consequences of privacy breaches listed were often centered around advertising, the main way to make money for WhatsApp parent company Meta. Participants talked about user profiling to maximize targeted ad earnings. *"And I realistically see the danger that the Meta Corporation uses that data to make money."* (P3). One participant based this model on Meta's reputation: *"And Meta is known for wanting to have, distribute and sell user data."* (P5), while another based it on Meta's monopoly position: *"Well I do know that data can be an insane currency. That's why I am always a little... because you can make a lot of money with it, especially when having a monopoly position it's somehow a resource and means."* (P4).

In this context, many participants expressed worry about the *creepy ads* phenomenon: *"a negative reaction to the impression that the marketer knows more about you than you want or expect them to know"* [8]. Advertisements seemingly follow a user around different devices or they receive advertisements on an item or product they did not consciously look up or search for on the internet but chatted about: *"When you actively text things about whales and then in your next web-browser you get advertisements for whales. And then you think hm that's weird."* (P1).

At the same time, some participants seemed to portrait the data collection and user profiling itself as the goal, perhaps as a way to more power and control.

> *"I am fully aware that individual information can be used there. And when Mark Zuckerberg says that with seven clicks on Facebook he knows me better than my closest relatives, then I also know [...] that a chat history that can be read, says a lot about me." —P10*

Attacks by Scammers. Some participants disclosed scammers as threats to security and privacy: *"There are these grandchildren tricks especially targeted at oldies. I am fully convinced that every human has these triggers that lead them to being careless or put common sense second. In that sense all of that is really dangerous."* (P10). Some participants also reported on their own experiences and situations where they themselves were targeted: *"What I thought was a little weird, this only happened last week, that I was texted by some random number on WhatsApp with a country code of +99 or something. They said "hello" and had a really odd profile picture. And I was confused and wondered where that had come from. That was really really weird."* (P1).

> *"You get so many spam messages somehow from people that aren't who they say they are. And I think that that would be my main concern." —P14*

Attacks by Law enforcement and intelligence agencies. Some participants assumed that government-run agencies eavesdrop on all WhatsApp messages: *"It's very clear to me that the world communication runs through America. When I do something, Homeland Security always listens in."* (P10) and *"Intelligence agencies, for example could definitely [eavesdrop]. I would say especially state resources."* (P14).

Attacks by Hackers and Criminals. One participant listed the threat of hackers eavesdropping on messages to later use private chat messages for extortion purposes. Users could suffer reputational damage from leaking of private photos: *"For photos I could imagine that there are problems. Especially for, let's say, spicier photos, that they get into the wrong hands and land on porn- or similar sites."*

(P6), or be endangered by being associated publicly to certain political viewpoints: *"If you think this further, political jokes of any kind could be used against you very quickly. I am very active in a protest-heavy leftist context and sometimes send protest pictures or calls to action and the likes. I think that could have negative consequences [...]."* (P4).

Attacks on End Device. Some participants noted that, if their end device was compromised, either remotely ("Hacked") *"But when someone hacks into my mobile phone, then they can already see all the messages because I can access them from my phone."* (P14). or physically

> *"What do I verify? If someone steals the SIM card from your phone and inserts it into a new one, do I verify you even though it's not you? In the end I still don't know who is holding the phone." —P11*

.

User Error. Some participants also mentioned user errors as privacy risks, like accidentally saving the wrong phone number for a contact, or accidentally sending a sensitive message to the wrong contact: *"Maybe I could have mistyped the number or something."* (P5).

WhatsApp Data Leaks. The risk of WhatsApp having a data leak or being hacked was touched on, as well. One participant stated: *"Well the things sit on a server somewhere for sure. Therefore, if there is a leak, I could imagine that somewhere some data may leak."* (P7).

WhatsApp gains rights to my data. Another worry, expressed by P3, was that WhatsApp would gain rights to all data they put into the system: *"I am under the impression that you hand over rights to all information that you enter to the app. Or, I don't know, when you send a photo, then you hand over rights to that photo."* (P3).

In summary, our participants associated WhatsApp's security with end-to-end encryption, which all were aware of and many could describe correctly. However, trust in the effectiveness and implementation of the encryption varied. While some participants expressed confidence in WhatsApp's security and trusted the company, many voiced skepticism——often based on gut feelings, Meta's reputation, or past scandals. Trust was often described as partial, conditional, or pragmatic, rooted more in WhatsApp's ubiquity than genuine confidence. Participants' threat models reflected this ambivalence, ranging from concerns about Meta's data practices to fears of surveillance, scams, hackers, or personal mistakes.

5.2 Key Transparency in WhatsApp

We report on the results of the second part of our interview, where participants interacted with the KT UI and queried about their perceptions and understanding of it.

5.2.1 Perceived Security of Key Transparency. After our participants completed the task (used the key transparency check), we asked how secure it made them feel. Explanations ranged from more secure, to no change, to less secure:

Did not feel secure / No change. One common theme was that the key transparency check did not feel secure and did not have an impact on the perceived security of WhatsApp: *"I think, if I wouldn't have done it, I would have felt the same."* (P5). When enquiring further, participants discussed several reasons for the position. One participant was skeptical of the system's speed: *"I found it to be maybe a bit too fast. [...] Just too fast for it to be more secure now."* (P2). Multiple participants remarked doubts that the check did anything at all: *"'It just does some animated thing, [...] it pretends to be quickly calculating something that... I doubt that it really did something.' —P3"*. Another reason for the lack of perceived security was the absence of action from the user side: *"This doesn't give me any security. They claim now that it would be secure. Anyone can claim stuff. [...] I didn't enter a password like for other encryptions or any other action where I would prove that I am me."* (P13). P1 expressed skepticism, because they expected having to go through a manual authentication ceremony to be secure: *"'I thought that if you want to encrypt it really, really hardcore, that you always have to do it in person.' —P1"*

Felt a bit more secure. Other participants reported feeling more secure after using the key transparency check. The level of increase in perceived security differed between participants, however. One participant stated: *"It would do something to me, it would give me a bit more security, but not enough that I would start sending passwords or credit card numbers or whatever."* (P2). A different participant reported a stronger increase in perceived security but questioned its usefulness: *"'It gives me a sense of security which, in reality, I don't need for my normal WhatsApp traffic.' —P10"*. The UI visuals were specifically mentioned to have increased P16's perceived security: *"It felt secure in so far as that you had to specifically press a button, then it showed something green, which obviously goes to show that it's totally super, then there was a lock, that is secure. And it verified that something worked."* (P16).

Felt less secure. Some interviewees expressed feeling less secure *after* using the key transparency check than before. Seeing a security option that the users had never seen before made them feel like they had missed out on securing their chats, sometimes raising concerns about what other security options they did not know about:

> *"When I would find out that this is a feature which existed for a long time but I never used it because I didn't know it and it increases my security, then I would feel less secure. Because I have the feeling that in all this mess of settings and whatever there are maybe more things that I should be doing to really be secure." —P11*

A similar sentiment was based on the perception that the key transparency check started the encryption and any messages before would be sent in clear text: *"I had the feeling this was always on. That you don't have to do this as well. I thought that it was end-to-end encrypted in every chat. So, no [it didn't make me feel secure], more the opposite."* (P6).

5.2.2 Key transparency confusion. When asked what they thought had just happened in the key transparency system, one participant replied: *"To be honest I can't imagine [what just happened], no. I don't know what exactly happened there."* (P5). For some other participants, this confusion posed a real problem: *"'Was automatically verified'. Does this mean anything now? [...] It doesn't explain itself to me.*

They write that they are now end-to-end encrypted, so what does this even change?" (P11). Not seeing a change in the system after using key transparency, one participant stated:

> *"Somehow I have the feeling that after doing this encryption, that you just do it for the sake of doing it and you don't really know what it even did. Okay you encrypted it now and does anything change? Not really. [...] That's a little, I find that a little, like, why? Why do I have to verify this?" —P1*

As outlined above, an average every day WhatsApp user can not be expected to understand the intricacies of a system like key transparency and it makes sense that after using it once, there would be some unanswered questions.

Repeat authentication for every message. Another common confusion regarding the key transparency system was about the needed frequency of use. Participants were unsure how often they had to re-do the automatic check: *"Would I have to do this more often, before every sent message? Or is it enough to do it once? Forever?"* (P6). Another participant remarked that they would not be willing to repeat the verification process regularly: *"But I don't think that, every time I text her, I would do this verification process again."* (P5). In theory, end-users only have to repeat the key authentication every time their chat partner's key material changes. This only occurs when registering or de-registering an account or device, or reinstalling the app. We further discuss this in the context of key change notifications in section 6.2.

5.2.3 Misconceptions. Our interviewees showed misconceptions about WhatsApp and key transparency, ranging from trivial to harmful. WhatsApp is closed source, so we assume their published white papers, blog articles, and customer support documentation as ground truth.

Encryption has to be turned on. One common misconception was that the E2EE would have to be manually turned on: *"When you start a new chat, there is this yellow box. And I mostly started ignoring it, but it always said something about end-to-end encryption and I think nowadays it says something about having to turn it on."* (P2). The yellow E2EE banner on top of every new chat reads: *Messages and calls are end-to-end encrypted. No one outside of this chat, not even WhatsApp, can read or listen to them. Tap to learn more.* Hence, E2EE is always on in every WhatsApp chat and this sentiment qualifies as a misconception. One dangerous false conclusion from this misconception is that messages were not encrypted before using the key transparency system. One participant thought that all previous messages were sent in clear text, leading to a worse perceived security overall.

Key transparency doesn't do anything. Another misconception was that key transparency is essentially just a front to increase end-user's trust and perceived security: *"I think it's just mimicry. [...] I wouldn't know what it supposed to have done now."* (P7). Another participant specified that no one could prove to them that the system actually has a use:

> *"It looked like something happened, whether something actually happened, how am I supposed to know? This doesn't tell me anything. No one showed me this, and I think no one can." —P9*

In summary, our participants' reactions to the KT UI were mixed: some felt slightly more secure, others felt no change, and a few felt less secure−−often based on confusion or misconceptions. Many struggled to understand what KT actually did, how often it needed to be used, or whether it actually had a function. Misconceptions included believing that encryption must be manually enabled by starting the KT check, or that key transparency was pure window dressing rather than a functioning security mechanism.

5.3 Sending Sensitive Data via WhatsApp

We asked interview participants whether they had ever *not* sent a piece of data or information via WhatsApp because it felt too sensitive for them. This question was included to pay attention to the *privacy paradox* — where actual end-user behavior can appear contradictory to their reported privacy concerns [2]. Pieces of information that were identified as sensitive by participants were: passwords, private conversations, private photos (of children), ID cards, health information, bank transfers, and other financial information.

*"I **would** use a different channel".* Some participants did not answer with a concrete instance where they didn't send something via WhatsApp because it was too sensitive, but still stated that they do not believe WhatsApp chats to be a secure place for any of their sensitive information: *"Some things I will definitely not put in there. [...] I don't send bank details, I don't send all those kinds of data via WhatsApp. Or passwords."* (P13). When further inquiring about specific separate channels that the participants would make use of, a set of different channels were mentioned.

First, phone calls were touched on as a potential alternative channel. However, different participants had different ideas of why a phone call would yield more security than WhatsApp. In some cases, the participant aimed to verify that Emma really is Emma, her number is correct and she really wants them to send her credit card information via WhatsApp:

> *"I would have called her first. And then I would have considered whether it was part of the conversation. If I know that we talked about this today, that she wants to book flights, then I would just send this. But if she hadn't said it... I would just call I think." —P14*

The main threat model here is that of spam or scam messages. The danger is not that of WhatsApp being inherently insecure in transmitting data, more that the participant may accidentally reveal the credit card information to the wrong person. The worry of WhatsApp not being secure enough to be used to send credit card information, however, was also mentioned as a reason to use a phone call: *"There are things that I would not trust WhatsApp with, that I rather sort out in person or on a phone call."* (P7). A final argument as to why to use a phone call instead was to separate the credit card number and the security code and expiry date onto different channels so that even if the credit card number were to fall into the wrong hands, it would be worthless without the rest of the information:

> *"This verification [code], I would send this separately or use a phone call. Whether this gives me better security, I don't know, but I would definitely... This credit card number is the permission to print money. And at least the security code I wouldn't want to send this way." —P10*

It is important to note that none of the participants mentioning phone calls specified whether they would call their chat partner on the mobile network or use an in-app WhatsApp phone call. Different phone call methods like mobile data, voice over IP (VoIP), platforms like Zoom or Skype, or in-app calls in WhatsApp or Signal have different security standards. Yet, when comparing the theoretical security level of most phone calls to WhatsApp chat messages, it is fair to say that, for most threat models, WhatsApp chats are held to higher security standards. This is due to the fact that WhatsApp utilizes the open-source Signal protocol.

Another option of transferring sensitive information was using email: *"[…] or I could send the number via email."* (P15). Out of the few participants that mentioned email, one correctly expressed their doubts that this change of channel increased the overall security, however: *"'Then there is also the classic email, but that's also not secure.' —P7"*

*"I **did** use a different channel".* In the previous quotes, participants stated that they *would* try to use a separate channel. What we were particularly interested in, however, to combat social desirability and self-reporting biases, was whether participants could actively remember that they *did* use a separate channel in the past to transmit data they deemed too sensitive for WhatsApp communication. This is what interview question 20, regarding privacy habits (*Have you ever not sent something via WhatsApp because it was too sensitive?*), was targeted at.

Again, phone calls were a popular choice for transmitting sensitive data. One participant reported: *"Back in the day there was a situation, I still remember this, where I was still living in a shared accommodation with someone and that person needed some card really urgently and asked me to send them a photo of it. I said 'no let's just have a quick phone call and I'll give you the information that way'."* (P4). Likewise, a separate participant stated to have rejected the convenience of sending information quickly via WhatsApp for the sake of alleged increased security:

> *"When I create a secure password somewhere it would be easier to send it via WhatsApp because they can get very complex those 18 or 16 digit things created by the password manager. […] When someone needs those passwords, […] I transmit them via phone call." —P7*

WhatsApp chats, depending on your threat model, usually offer better security properties than phone calls. Thus, the switch from WhatsApp to a phone call may not increase security at all and, depending on the kind of phone call, may actually decrease it.

Email was also reported to have been used as an alternative channel in the past: *"I know that at some point back in the day I had to send a scan of my ID card and I consciously did not use WhatsApp for that. […] I actually think that I sent it as an email attachment which I did because I knew that it wasn't some shady email provider. I think it was GMX or something."* (P16). Interestingly, this participant actively thought about the email provider which, in normal email traffic, could be a potential attacker.

Some noteworthy edge cases that arose during the interviews were encrypted cloud sharing: *"Nowadays I work with Tresorit[3]. That's a provider for secure data exchange."* (P10), Signal: *"But also after a weekend trip when you took 30 pictures or something and you send them to your friends, I have done this not via WhatsApp as well. […] For example on Signal."* (P5), and password managers: *"My partner and I send each other passwords via password-safes if we both need them."* (P6).

"I send sensitive data via WhatsApp". Other participants had less reluctance and said that they would generally be willing or even reported on cases where they shared sensitive information on WhatsApp. One position was that taking the small risk of something happening to the data would be acceptable: *"I think I use WhatsApp consciously enough that I accept that there may be a small security risk and then I just accept it for what it is."* (P3). Another participant stated that they share sensitive information on WhatsApp against their better judgment: *"I think usually I would just send a photo, to be honest. It's probably not the most secure solution but I'll just do it anyways."* (P6). Similarly, the convenience and usability of WhatsApp was mentioned to outweigh the security risks it may bring:

> *"I think it would be most convenient to just send it via WhatsApp. I think there are other methods but they just take way more time and are a lot more complex." —P5*

In summary, many participants identified various types of information they considered too sensitive for WhatsApp, such as passwords or financial data. While some could recall specific instances where they used alternative channels like phone calls, email, or secure cloud services, others resorted to reporting that they *would* use other channels than WhatsApp for sensitive information. Some participants reported that they are okay with sending sensitive data via WhatsApp−−due to convenience, habit, or perceived acceptability of small risks.

6 Discussion

We discuss our findings on user perceptions of WhatsApp's security, perceptions of key transparency, concluding recommendations for both industry and academia. We also draw attention to issues highlighted by our study result that require attention when designing future KT systems for humans.

6.1 Perceptions of WhatsApp's Security

Our results cover a broad spectrum of user viewpoints, ranging from no trust in WhatsApp to high trust in WhatsApp.

Applying the theoretical framework of unidirectional trust (cf. subsection 2.3) to the instance at hand, the WhatsApp user (trustor) puts their trust into the WhatsApp company (trustee), as the service operator. By sending sensitive information and generally using WhatsApp, an end-user is willing to be vulnerable to potential attacks by WhatsApp to misuse the end-user's private messages, or skimp on security. When inquired about the source of their trust

[3] https://tresorit.com

or mistrust, participant reasonings can be dissected into the three trust components (cf. subsection 2.3), i.e. participant's assessments of WhatsApp's *capability*, *benevolence*, and *integrity*. Our findings imply that users generally do not doubt WhatsApp's *capability*, but often question their *benevolence* or *integrity*.

As all three trust components are relevant for users' ultimate trust decisions, chat app vendors should pay special attention to these imbalances when designing public communications as well as in the app's UI. Trust-building could be supported by UI elements that reinforce not only technical security but also company accountability, transparent monetization strategies, and user agency.

Our findings contrast those of Dechand et al.[17], who found almost no awareness for end-to-end encryption in their interviews in 2017, in that our participants directly mention E2EE as a, perhaps *the*, security feature of WhatsApp. We accredit this to WhatsApp's efforts in tirelessly communicating about E2EE to their users in ad campaigns [15, 31] and more importantly in WhatsApp's user interface, with hints at the beginning of every new chat over the past 7+ years since those interviews were conducted, as in Figure 2. The fact that all our participants knew the term end-to-end encryption, but only a fraction of them believed that it effectively keeps WhatsApp from reading message contents, reveals the sincere limitation of even these large-scale communication efforts, but also shows that they are not in vain.

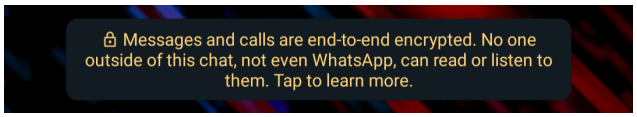


Figure 2: Info box about end-to-end encryption, displayed at the top of every new chat in WhatsApp—a part of WhatsApp's long ongoing, and partly successful efforts to communicate about E2EE.

6.2 Perceptions of Key Transparency

Our interviewees reacted in different ways when exposed to and inquired about WhatsApp's key transparency UI. While some were quick to disregard the KT page as window dressing, a few actually reported feeling more safe using WhatsApp after performing the KT check. A recurring sentiment was confusion over the necessity of user-initiated verification steps when encryption was presumed to function automatically. However, more worryingly, some participants felt *less* secure about the confidentiality of WhatsApp after using KT, either due to a) a misunderstanding that KT check is needed to turn on encryption or b) a previous overestimation of message confidentiality, assuming there was no possible way WhatsApp could read messages. We want users to makes informed decisions about their chat app usage regarding security and privacy. In the experiences our interviewees shared with us, they however often switched to—depending on ones threat model—less secure channels like email or phone calls when they deemed WhatsApp as inappropriate for sending sensitive information. Finding a digestible way to convey threat models for different communication channels at once is an unsolved design challenge worth investing further.

A Nuanced Shift in Threat Models. The majority of our interviewees did not show a change of security perception regarding WhatsApp after conducting a key transparency check for the first time. We initially hoped that key transparency would make users feel more secure when using WhatsApp. After shedding more light on end-user's practical threat models in our interviews, it seems reasonable and understandable that interacting with WhatsApp's KT UI would not have a big impact:

While the security gain (due increased amounts of effort required for attacks on end-to-end encryption) might seem clear for security researchers, certain threats remain, so to everyday users, it might simply occur that *they have to trust WhatsApp*, no matter what. To fully understand the nuanced security benefits of KT, users would require structural knowledge of chat app infrastructure and encryption protocols (cf. subsubsection 2.2.1).

Key Transparency Confusion and Misconceptions. We find that the current KT UI, without further explanatory material, can lead to misconceptions and confusion, potentially decreasing perceived security users have of WhatsApp. We believe the most promising way to reduce those problems is to ease the burden put onto users, by working towards automating KT checks fully and only show well designed, actionable warnings to users when an attack has been detected.

Key Change Notifications. Another noteworthy issue that arose in our interviews was the fact that it was unclear to users how often, or at what times, performing a key transparency check would be sensible. Starting a key transparency check would be sensible every time a chat partner changed their encryption keys, since a key change might result from a MitM attack mounted by the server operator. This, however, is never explained in the app and participants are rightly confused: In the default configuration, the current version of WhatsApp does not even show notifications for key changes. To receive any indication of when a key transparency check might have to be repeated for a contact, users have to manually enable "Security Notifications" in WhatsApp's settings. If enabled, when a contact's encryption key changes, a security notification, as seen in Figure 3, is shown as part of the corresponding chat history. If these notifications are not enabled, users might feel secure after verifying their contacts encryption key, but would not be notified if the keys changed because of a server-operator mounted MitM attack, rendering the manual checks much less useful.

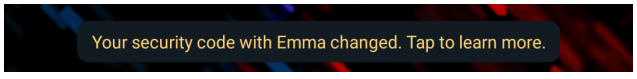


Figure 3: Security notification in WhatsApp displayed in the chat when a contact's encryption key changes. These notifications are disabled by default, but can be turned on in WhatsApp settings.

6.3 Negative Consequences of Key Transparency

While previous research shows that there are no theoretical security downsides to implementing KT, our study unveils issues that require attention when designing future KT systems for humans.

Underconfidence. When our participants were exposed to the KT UI, some thought the key transparency check did not *verify* the E2EE, but instead *started* the encryption, and therefore concluded that for every chat were they don't run a KT check, messages are transmitted unencrypted. This could ultimately lead to users feeling less secure when using WhatsApp and possibly switching to a different "secure" messaging app, which does not offer KT checks, which could be *less* secure.

Overconfidence. At the same time, other participants showed increased trust in WhatsApp's security after running a KT check. In combination with WhatsApp's persuasive claims that "No one outside of this chat, not even WhatsApp" could read the messages, a successful KT check could lead to overconfidence, unaware of persisting threats to message confidentiality like end-user device threats, cloud backups, and client-side code threats (cf. subsubsection 2.1.1).

Reliance on 3rd Party Witnesses. Some participants noted that they were unsure about what to do if they ran a KT check and it would not succeed. The KT system Parakeet[29], and thus WhatsApp's implementation of KT, relies on a trusted third party service ("Witness") to perform auditing of the KT Auditable Key Directory, and the KT Signing Service to prevent equivocation attacks (i.e. prevent WhatsApp serving different keys for the same identity to different users). This reliance means that, if for some reason, the 3rd party service is unavailable, all KT checks fail. Failing KT checks could seriously impact the trust relationship between WhatsApp and it's users, even though the availability of the 3rd party service it outside of WhatsApp's control.

7 Outlook

WhatsApp offering authentication ceremonies is important for message privacy, even if almost no one uses them. Having just the possibility to be checked is a *credible threat* (a term borrowed from game theory) and thus has a valuable *deterrence effect*: Because WhatsApp server operators could be checked at any time, they are incentivized to behave honestly and not mount MitM attacks. The implementation of more convenient verification, in the form of KT increases this effect by making checks more accessible. We note that to this day, there is no evidence that WhatsApp has ever tried to mount a MitM attack on their users.

We found that some users might perceive WhatsApp as less secure after learning about KT and the threat it protects from. WhatsApp might not want to explain the threat model of MitM-Attacks in detail since it can make users feel less secure—and prompt a switch to a competing chat app which does not mention any possibilities of attacks. The most promising solution to this conundrum is to fully automate key transparency. This would simplify the threat model aspect to "WhatsApp can't read your messages", or at least "WhatsApp server operators can't read your messages", since the client code, written by developers, still needs to be trusted. If WhatsApp wants to truthfully claim that no one, not even WhatsApp, can read their users' messages, it is imperative for them to secure key authenticity, or at least key consistency, in a more automated matter.

7.1 Future Work

We encourage future research to replicate and expand on our findings using larger and more representative samples, including participants with lower digital literacy, non-German cultural backgrounds, or investigating different implementations of key transparency. Additionally, our findings draw attention to exciting new research avenues:

Key History Checks. In its current implementation, key transparency in WhatsApp does not offer key history checks. Key history checks, as proposed in *Parakeet* [29], would allow a user to not only check whether they are *currently* a victim of a MitM attack, but also whether there were any suspicious changes to the key material distributed for their identity in the past. Key history check would need an additional user interface, in which users get a concise overview about all (more or less recent, due to KT compaction) key change events associated with their identity. Designing such an interface in an comprehensible manner, which would avoid misinterpretation of key changes (and thus lead to mistrust), is a challenging task which would benefit greatly from future research and user studies.

KT Failure Warnings. In our study, people indicated that they would not know what to do if the key transparency process fails. Creating some actionable advice for these very rare warnings would likely be helpful. Akin to the progress made on TLS warnings, the focus should be on avoiding false-positives, nudging users towards safe behavior, and creating actionable warning text. This is tricky, because what users should do if the key verification fails is not obvious.

Explaining KT and Chat App Threat Modeling. For better ecological validity of our interview study results, we intentionally did *not* explain KT to our participants before having them use and judge it. We hypothesize that, if simple exposure to the KT UI will, as shown in our interviews, not meaningfully increase perceived security and trust, maybe a better understanding of KT would. If understood, the ability to monitor the trustee, e.g. via Key Transparency, should decrease the need for trust and thus increase the perceived security of a system. At the same time, the fact that WhatsApp voluntarily implements such a monitoring system, might deem it more trustworthy in the eyes of its users. Future research might explain KT to users and measure if understanding KT and its nuanced threat model can increase trust in or perceived security of chat apps. Research like this would inform science outreach efforts to teach broader populations about the differences in security across "secure" chat apps. If we are not able to convey fundamental threat models in digestible ways, end-users are left vulnerable to battling and possibly misleading marketing claims.

7.2 Recommendations for Practitioners

We summarize the recommendations from our discussion, that may support industry practitioners involved in the design of secure chat apps and of communication efforts related to chat app security.

(1) Better address the trust components *benevolence* and *integrity* in public communication as well as in the in-app UI. KT UI copy should not only highlight technical security, but also communicate the company's values and practices clearly to foster overall trust.
(2) Clarify the purpose and security benefit of key transparency in-app to avoid confusion. Our participants were confused about what the KT check does, when it's needed, and why it matters. Work towards a UI and documentation that can briefly and clearly explain this.
(3) Automate key transparency checks where possible to reduce the need for user action and instead surface *actionable* warnings only when potential attacks are detected. Keep iterating on it to, in the future, achieve KT's initial design goal: Increased security by regular key consistency monitoring with minimal user input, by fully automating the process and implementing key history checks.
(4) Integrate a link to enable "Security Notifications" directly in the KT UI and explain why they are needed. In-app nudges and contextual links could encourage more meaningful KT use. We understand that enabling "Security Notifications" by default, while great for security, might not be immediately desirable, since it can scare uninformed users into using other channels, as evident from our interviews.
(5) Work towards conveying chat app threat models in digestible formats. Our study confirms findings from prior work that users have misconceptions about the security of different communication channels (E-Mail, SMS, different Chat App). We want users to use the most appropriate tool for their privacy needs.

As security researchers, we commend WhatsApp for pioneering large-scale key transparency. This large-scale implementation sets an important precedent, offering the potential to, in future iterations, enhance security for all users by further deterring service operators from mounting machine-in-the-middle attacks with fully automated key transparency checks.

Acknowledgments

We thank the anonymous reviewers and our revision editor for their valuable feedback. We are sincerely grateful to all participants who took part in our study. We also thank Kevin Lewi for answering our questions about how Key Transparency works in WhatsApp. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2092 CASA - 390781972.

AI Acknowledgement

The authors used ChatGPT-4o to revise text drafts throughout all sections of the paper to correct any typos, grammatical errors, and awkward phrasing. No AI outputs were copied into the manuscript without thoroughly checking that the initial text's content and meaning were unaltered. In almost all cases, the authors made further adjustments to better match the author's writing style and intentions.

References

[1] Ruba Abu-Salma, M Angela Sasse, Joseph Bonneau, Anastasia Danilova, Alena Naiakshina, and Matthew Smith. 2017. Obstacles to the adoption of secure communication tools. In *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, Institute of Electrical and Electronics Engineers, San Jose, CA, USA, 137–153.
[2] Alessandro Acquisti. 2004. Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM conference on Electronic commerce*. Association for Computing Machinery, New York, NY, USA, 21–29.
[3] Mashari Alatawi and Nitesh Saxena. 2023. SoK: An Analysis of End-to-End Encryption and Authentication Ceremonies in Secure Messaging Systems. In *Proceedings of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. Association for Computing Machinery, New York, NY, USA, 187–201.
[4] Joël Alwen, Sandro Coretti, and Yevgeniy Dodis. 2019. The double ratchet: security notions, proofs, and modularization for the signal protocol. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, Darmstadt, Germany, 129–158.
[5] 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 *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. Association for Computing Machinery, New York, NY, USA, 25–36.
[6] Matilda Backendal, Hannah Davis, Felix Günther, Miro Haller, and Kenneth G Paterson. 2024. A formal treatment of end-to-end encrypted cloud storage. In *Annual International Cryptology Conference*. Springer, Santa Barbara, CA, USA, 40–74.
[7] Alexander Bienstock, Jaiden Fairoze, Sanjam Garg, Pratyay Mukherjee, and Srinivasan Raghuraman. 2022. A more complete analysis of the signal double ratchet algorithm. In *Annual International Cryptology Conference*. Springer, Santa Barbara, CA, USA, 784–813.
[8] Jason Bloomberg. 2014. *The Big Data Marketing Creepiness Factor*. Intellyx. https://intellyx.com/2014/11/10/the-big-data-marketing-creepiness-factor/.
[9] Kathy Charmaz. 2014. *Constructing Grounded Theory*. Sage, Thousand Oaks, CA, USA.
[10] Melissa Chase, Apoorvaa Deshpande, Esha Ghosh, and Harjasleen Malvai. 2019. Seemless: Secure end-to-end encrypted messaging with less trust. In *Proceedings of the 2019 ACM SIGSAC conference on computer and communications security*. International Association for Cryptologic Research (IACR), Santa Barbara, CA, USA, 1639–1656.
[11] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* 3 (2015), 222–248.
[12] Katriel Cohn-Gordon, Cas Cremers, Benjamin Dowling, Luke Garratt, and Douglas Stebila. 2020. A formal security analysis of the signal messaging protocol. *Journal of Cryptology* 33 (2020), 1914–1983.
[13] Daniel Collins, Doreen Riepel, and Si An Oliver Tran. 2024. On the Tight Security of the Double Ratchet. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*. Association for Computing Machinery, Copenhagen, Denmark, 4747–4761.
[14] Juliet Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift für Soziologie* 19, 6 (1990), 418–427.
[15] Basudha Das. 2022. WhatsApp issues eye-catching multi-layer lock privacy ad. https://www.businesstoday.in/technology/news/story/whatsapp-issues-eye-catching-multi-layer-lock-privacy-ad-take-a-look-347122-2022-09-13 [Online News Article; accessed 2025-02-20].
[16] Gareth T Davies, Sebastian Faller, Kai Gellert, Tobias Handirk, Julia Hesse, Máté Horváth, and Tibor Jager. 2023. Security analysis of the whatsapp end-to-end encrypted backup protocol. In *Annual International Cryptology Conference*. Springer, The Hague, The Netherlands, 330–361.
[17] Sergej Dechand, Alena Naiakshina, Anastasia Danilova, and Matthew Smith. 2019. In encryption we don't trust: The effect of end-to-end encryption to the masses on user perception. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, Stockholm, Sweden, 401–415.
[18] Verena Distler, Tamara Gutfleisch, Carine Lallemand, Gabriele Lenzini, and Vincent Koenig. 2022. Complex, but in a good way? How to represent encryption to non-experts through text and visuals–Evidence from expert co-creation and a vignette experiment. *Computers in Human Behavior Reports* 5 (2022), 100161.
[19] Verena Distler, Carine Lallemand, and Vincent Koenig. 2020. Making encryption feel secure: Investigating how descriptions of encryption impact perceived security. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, Genova, Italy, 220–229.
[20] Matthias Fassl and Katharina Krombholz. 2023. Why I Can't Authenticate—Understanding the Low Adoption of Authentication Ceremonies with Autoethnography. In *Proceedings of the 2023 CHI Conference on Human Factors in*

Computing Systems. Association for Computing Machinery, Hamburg, Germany, 1–15.
[21] Tilman Frosch, Christian Mainka, Christoph Bader, Florian Bergsma, Jörg Schwenk, and Thorsten Holz. 2016. How secure is TextSecure?. In *2016 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, Springer, Berlin, Germany, 457–472.
[22] Nina Gerber, Verena Zimmermann, Birgit Henhapl, Sinem Emeröz, and Melanie Volkamer. 2018. Finally johnny can encrypt: But does this make him feel more secure?. In *Proceedings of the 13th International Conference on Availability, Reliability and Security*. Association for Computing Machinery, Hamburg, Germany, 1–10.
[23] Amir Herzberg and Hemi Leibowitz. 2016. Can Johnny finally encrypt? Evaluating E2E-encryption in popular IM applications. In *Proceedings of the 6th Workshop on Socio-Technical Aspects in Security and Trust*. Association for Computing Machinery, Trento, Italy, 17–28.
[24] Amir Herzberg, Hemi Leibowitz, Kent Seamons, Elham Vaziripour, Justin Wu, and Daniel Zappala. 2020. Secure messaging authentication ceremonies are broken. *IEEE Security & Privacy* 19, 2 (2020), 29–37.
[25] Erin Kenneally and David Dittrich. 2012. *The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research*. Technical Report. Homeland Security. https://doi.org/10.2139/ssrn.2445102
[26] Ben Laurie. 2014. Certificate transparency. *Commun. ACM* 57, 10 (2014), 40–46.
[27] Sean Lawlor and Kevin Lew. 2023. Deploying key transparency at WhatsApp. https://engineering.fb.com/2023/04/13/security/whatsapp-key-transparency/.
[28] Julia Len, Melissa Chase, Esha Ghosh, Kim Laine, and Radames Cruz Moreno. 2023. OPTIKS: An Optimized Key Transparency System.
[29] Harjasleen Malvai, Lefteris Kokoris-Kogias, Alberto Sonnino, Esha Ghosh, Ercan Oztürk, Kevin Lewi, and Sean Lawlor. 2023. Parakeet: Practical key transparency for end-to-end encrypted messaging. Cryptology ePrint Archive, Paper 2023/081.
[30] Moxie Marlinspike. 2016. WhatsApp's Signal Protocol integration is now complete. https://signal.org/blog/whatsapp-complete/.
[31] Jim Martin. 2024. New WhatsApp ads promote end-to-end encryption - Tech Advisor. https://www.techadvisor.com/article/743988/new-whatsapp-online-tv-radio-ads-promote-end-to-end-encryption.html [Online News Article; accessed 2025-03-01].
[32] Roger C Mayer, James H Davis, and F David Schoorman. 1995. An integrative model of organizational trust. *Academy of management review* 20, 3 (1995), 709–734.
[33] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (Nov. 2019), 23 pages. https://doi.org/10.1145/3359174
[34] Marcela S Melara, Aaron Blankstein, Joseph Bonneau, Edward W Felten, and Michael J Freedman. 2015. {CONIKS}: Bringing key transparency to end users. In *24th USENIX Security Symposium (USENIX Security 15)*. USENIX Association, Washington, D.C., USA, 383–398.
[35] Thibault Meunier and Mari Galicer. 2024. Cloudflare helps verify the security of end-to-end encrypted messages by auditing key transparency for WhatsApp. https://blog.cloudflare.com/key-transparency/.
[36] Andrew J. Paverd and Andrew C. Martin. 2014. Modelling and Automatically Analysing Privacy Properties for Honest-but-Curious Adversaries. https://api.semanticscholar.org/CorpusID:211141069
[37] Trevor Perrin and Moxie Marlinspike. 2016. The double ratchet algorithm.
[38] Svenja Schröder, Markus Huber, David Wind, and Christoph Rottermanner. 2016. When SIGNAL hits the fan: On the usability and security of state-of-the-art secure mobile messaging. In *European Workshop on Usable Security. IEEE*. Internet Society, San Diego, CA, USA, 1–7.
[39] We Are Social, DataReportal, and Meltwater. 2024. Most popular global mobile messenger apps as of April 2024, based on number of monthly active users (in millions). https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/.
[40] Ryan Stedman, Kayo Yoshida, and Ian Goldberg. 2008. A user study of off-the-record messaging. In *Proceedings of the 4th Symposium on Usable Privacy and Security*. Association for Computing Machinery, New York, NY, USA, 95–104.
[41] Christian Stransky, Dominik Wermke, Johanna Schrader, Nicolas Huaman, Yasemin Acar, Anna Lena Fehlhaber, Miranda Wei, Blase Ur, and Sascha Fahl. 2021. On the Limited Impact of Visualizing Encryption: Perceptions of {E2E} Messaging Security. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. USENIX Association, Virtual Conference, 437–454.
[42] Anselm Strauss and Juliet M Corbin. 1997. *Grounded Theory in Practice*. Sage, Thousand Oaks, CA, USA. 288 pages.
[43] Apple Support. 2025. iCloud data security overview - Apple Support. https://support.apple.com/en-us/102651 [Online; accessed 2025-02-27].
[44] Nik Unger, Sergej Dechand, Joseph Bonneau, Sascha Fahl, Henning Perl, Ian Goldberg, and Matthew Smith. 2015. SoK: secure messaging. In *2015 IEEE Symposium on Security and Privacy*. IEEE, San Jose, CA, USA, 232–249.
[45] 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 *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. USENIX Association, Baltimore, MD, USA, 47–62.
[46] Elham Vaziripour, Justin Wu, Mark O'Neill, 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 *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association, Santa Clara, CA, USA, 29–47.
[47] WhatsApp. 2023. *Key Transparency Overview (Technical Whitepaper)*. Technical Report. WhatsApp. https://www.whatsapp.com/security/WhatsApp-Key-Transparency-Whitepaper.pdf
[48] WhatsApp. 2024. *Encryption Overview*. Technical Report. WhatsApp.
[49] Justin Wu, Cyrus Gattrell, Devon Howard, Jake Tyler, Elham Vaziripour, Daniel Zappala, and Kent Seamons. 2019. "Something isn't secure, but I'm not sure how that translates into a problem": Promoting autonomy by designing for understanding in Signal. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. USENIX Association, Santa Clara, CA, USA, 137–153.

A Pre-Interview Screening Questions

The short screener was linked in the study invitations. The collected data helped us recruit a diverse sample with variation in demographics, occupations, and educational backgrounds.

(1) Is your primary mobile phone an Android phone? (No → stop)
(2) Is the main focus of your studies or job on IT security? (Yes → stop)
(3) What is your gender? (Woman | Man | Non-binary | Prefer to self-describe)
(4) What is your age? (blank)
(5) What is the highest degree or level of school you have completed? (Lower School Leaving Certificate | General University Entrance Qualification (Highschool) | Bachelor's Degree | Master's Degree | State Examination | Ph.D | Other)
(6) What is your current occupation? (blank)
(7) Please indicate how much you agree with each statement below (5-point Likert Scale)
   - I have a good understanding of smartphones and the internet.
   - I often ask other people for help when I am having problems with my smartphone.
   - I am often asked for help when other people have problems with their smartphone.

B Complete Interview Guide

This interview guide was used for conducting the semi-open interviews.

Briefing

Introduce yourself and thank the participant for taking their time to take part in the study. Together with the participant, go through the informed consent form.

Did you bring your Android phone and does it have enough battery for the study?

Do you have any questions for me at this point before we start?

Pre-Task Interview

Start the audio recording and let the participant know that they are now being recorded.

(1) How much do you use chat apps on a regular day?
(2) What chat apps do you use?

(3) When did you start using chat apps?
(4) (How has your usage of chat apps changed over the years?)

(5) In general... is WhatsApp secure? (Please explain.)
(6) Imagine you send someone a message on WhatsApp. What could go wrong from a security or privacy perspective?
   a. Can someone eavesdrop? Who could eavesdrop?
   b. Are there any countermeasures to eavesdroppers?
   c. Have you heard of the term end-to-end encryption?
   d. Do you think WhatsApp messages are encrypted? If so, could anyone still eavesdrop? If yes, who?

In the following questions, when I say "WhatsApp", I am referring to the company behind the app, not the app itself.

(7) How would you assess WhatsApp's ability to protect your chats and messages?
(8) What do you think is WhatsApp's motivation behind protecting chats and messages in general?
(9) How would you assess your trust in WhatsApp? Please explain.
   a. *If the participant speaks of low trust*: Why do you use the app despite not trusting WhatsApp?

Task 1 - Key Transparency

Read/Explain the scenario and the task description to the participant.

Imagine you meet your best friend Emma for a coffee. Emma tells you that she just got a new SIM card and had to change her number. She gives you a little slip of paper with her number on it so you can conveniently add her as a new contact to your phone.

TASK: Add Emma to your contacts.

You catch up for a while and eventually Emma leaves for a work appointment. You stay for a few more minutes and see that she forgot her bag under the table. You take it with you and send her a message, saying that you plan on bringing it to her house the next day.

TASK: Send Emma a WhatsApp message saying that you have her bag and can bring it to her tomorrow.

Respond in Whatsapp Web as Emma with the message: "Hey, thank you for taking my bag! Could you do me a huge favor and send me the number of the black credit card in my wallet? I really need to book these flights tonight."

TASK: Send Emma her credit card number. Be sure to verify the encryption in your chat with Emma ***before*** *sending her the credit card information. You can automatically verify the encryption by tapping on Emma's name at the top of the chat window and selecting the "Encryption" submenu.*

Interview - Key Transparency

If the participant did not consent to showing their screen, ask them to describe what they saw during the task to make sure they were in the correct setting.

(10) Please describe in your own words what just happened.
   a. What do you think happened inside the system?
(11) (On a scale of one to ten), how secure or insecure did this experience feel to you? Please explain.
(12) How did what you just saw change your trust in the Whatsapp company? Please explain.
(13) How did you generally like the procedure of verifying the encryption?
   a. What did you like best?
   b. What would you most want to change?
(14) (Did you run into any problems or difficulties while verifying the encryption? If so, please describe them.)

Final Questions

(20) Have you ever *not* sent something via WhatsApp because it was too sensitive?
   a. What kind of data was too sensitive for you to send it via WhatsApp?
   b. How did you send the data instead?
(21) *Show beige E2EE-banner on top of new chat to the participant.*
   a. What does this tell you? Do you benefit from this?
   b. Have you ever questioned this?
   c. Does WhatsApp do what they say here?
   d. Does WhatsApp have integrity?
(22) The process of scanning a QR code is also called an "authentication ceremony".
   a. Did you know this or a similar system before this study?
   b. Have you ever used the authentication ceremony in WhatsApp or any other app? If so, please describe the app and the process.

C Interview Study Codebook

We present our resulting codebook resulting from the merging process described in Section 4.1:

Merged Codebook

1. Chat App Usage
1.1 Reasons
2. Usability
2.1 Key Transparency
2.2 QR Code
3. Perceived Security
3.1 Is WhatsApp secure?
3.2 Threat Models / What can go wrong?
3.3 Threat Actors
3.4 Key Transparency
3.5 QR Code
3.6 Extra work makes me feel more secure
3.7 Sending sensitive data via WhatsApp
4. Trust
4.1 Do you trust WhatsApp?
4.2 Is WhatsApp capable of securing my data?
4.3 Is WhatsApp motivated to secure my data?
4.4 Changed trust in WhatsApp?
4.5 WhatsApp's Integrity to secure my data
5. "I don't know" / Uncertainty
6. Misconceptions / Confusion about KT and QR
6.1 Misconceptions
6.2 KT Confusion
6.3 QR Confusion
7. Previous Knowledge
7.1 E2EE

7.2 Key Transparency
7.3 QR-Code
7.4 Yellow E2EE Intro Banner
7.5 Key Change Security Notifications
7.6 Valid Assessment of KT's / QR's functionality
8. Sources of Information
8.1 Media
8.2 Yellow E2EE Banner
8.3 My friend
8.4 My tech-savvy friend
9. Future Use
9.1 Intent: Learn about the system
9.2 Intent: I won't authenticate keys in the future
9.3 Intent: Authenticate keys more in the future
9.4 Intent: Use WhatsApp less in the future
10. Other
10.1 Types of personal information
10.2 Key Transparency adds value / no value
10.3 They can have my data / "I ain't no king"
10.4 US laws are not as strict as EU laws
10.5 Someone else takes care of my security
10.6 I think I won't understand the WhatsApp info texts
11. Good Quotes