How to Explain Trusted Execution Environments (TEEs)?

Carolina Carreira
Carnegie Mellon University
IST University of Lisbon, INESC-ID

McKenna McCall
Carnegie Mellon University

Lorrie Faith Cranor
Carnegie Mellon University

1 Introduction

Trusted Execution Environments (TEEs) are isolated environments for executing code that guarantee the authenticity of the executed code, the integrity of the runtime states, and the confidentiality of its code and data [7]. Previous work investigates how the presence of TEEs effects privacy norms for smart home devices [6]. While TEEs can fill an important gap in system security, without clear and accessible explanations of TEEs and the guarantees they offer, they may do little to address users' *perceptions* of safety. Indeed, explaining security concepts like TEEs is important as it can empower users to make informed choices with technology [2, 4, 5].

In this work-in-progress study, we investigate potential TEE explanations to enhance both *understanding* of the capabilities that a TEE does (and does not) have and *trust* in TEE-enhanced technologies in the context of specific scenarios by answering the following research questions:
RQ1: What components might be in a TEE explanation?
RQ2: Which components should be included to improve TEE understandability? Which components should be included to improve trust in the TEE-enhanced technology?
RQ3: Is there an overall best explanation? Or do different technologies benefit from different TEE explanations, like home IoT and medical research applications?
RQ4: Does an FAQ improve understandability and/or trust?
RQ5: Which aspects of the scenario do participants report contributing to the belief that their data would be safe/unsafe?

To the best of our knowledge, our study is the first to focus on develping accessible TEE explanations.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2024.
August 11–13, 2024, Philadelphia, PA, United States.

2 Methods

We built candidate TEE explanations based on existing explanations found in the wild and evaluated them on their ability to enhance comprehension and trust via an online survey.

Identifying Candidate TEE Explanations To write our candidate TEE explanations, we used a methodology similar to prior work that developed explanations for differential privacy [3] to collect diverse in-the-wild TEE explanations (**RQ1**). We conducted a Google search using the term "Trusted Execution Environment" and restricted the results to the last five years. We retrieved 42 explanations from the first four pages of results spanning diverse sources, such as news articles and reports. Eight additional explanations were obtained through searches targeting well-known, general audience platforms like the New York Times, Intel, and Forbes. The aim was to supplement our collection with explanations from popular resources known for their broad readership.

From the initial 50 explanations, we discovered 12 did not have substantive TEE explanations. We removed 2 other explanations because they were incorrect/misleading. We analyzed the remaining 36 explanations to identify themes (i.e., components) to test in our experiments.

Two of the authors independently reviewed the explanations to identify additional themes, where each explanation might be assigned multiple themes. The coding process began with a small number of initial codes based on the authors' prior knowledge of TEEs. After reviewing all explanations, the coders discussed the themes to develop a shared codebook. They repeated the process of reviewing explanations, coding, discussing all disagreements, and refining the codebook twice more until 100% agreement was reached during the last meeting. The final codebook can be found in Appendix A.

Evaluating Candidate TEE Explanations To evaluate our candidate TEE explanations, we conducted an online survey. The purpose of the survey was to evaluate our candidate TEE explanations to identify which were best at enhancing understanding and trust (**RQ2**). Participants were asked to imagine

themselves in two scenarios (**RQ3**), one where they were invited to participate in a medical research study and another where they are shopping for a smart home device. In both scenarios, they are told that their data is stored in the cloud and that the data and computations on the data are protected by a TEE. They are given one of our candidate TEE explanations to read after the scenario is introduced.

We evaluated understandability via 10 True/False questions (these questions can be found in Appendix B). We assessed trust by asking participants to rate their willingness to participate in the medical research study (for the medical research scenario), willingness to purchase the smart home device (for the smart home device scenario), and how safe they believe their data would be, each on a 3-point Likert scale. We analyzed the True/False statements via logistic regressions and the willingness and safety questions using ordinal logistic regressions. At the end of each scenario, we gave participants the opportunity to ask us any lingering questions they have about TEEs. We applied similar techniques to analyze the questions as for the initial TEE explanations found in the wild, described above.

We recruited 469 adults located in the US who are fluent in English through Prolific using quotas [1] to ensure approximately equal numbers of men and women. Participants were paid $2.50 (median completion time approx. 10 minutes). We refined the survey questions through multiple pilots. The survey and consent form were approved by CMU's IRB.

3 TEE Explanations and Preliminary Results

We focus on the development of our candidate TEE explanations (**RQ1**) and brief, preliminary results of our evaluation (**RQ2-3**). We plan to address **RQ4-5** in a follow-up study.

Candidate TEE Explanations Our finalized codebook had 14 codes, the most common being *Confidentiality*, *Isolation*, and *Hardware*. We build all explanations around three fundamental concepts: *Confidentiality*, *Isolation* (because they are two of the most common themes), and *Integrity* (since it is dual to confidentiality). All explanations involve these themes, so they were not evaluated in our experiments. To keep the number of treatments tested in our surveys reasonable, we also did not test the least common themes, *Reputation*, *Verified*, and *Attestation*, which appeared in 5 or fewer explanations in the wild. While writing candidate explanations, we also eliminated *Threat* and *Techniques* because the resulting explanations were similar to other themes.

As a result, we had six themes to evaluate in our experiments. The structure of an explanation is: (1) a high-level sentence introducing the concept of a TEE as a security mechanism (themes: *Hardware*, *Trust*, or *Unsubstantial*), followed by (2) a sentence introducing the concepts of isolation, confidentiality, and integrity (themes: *Technical* or *Non-Technical*), and, only for some explanations, (3) a third sentence introducing a specific threat that a TEE can prevent (theme: *Prevents* or no third sentence). Our candidate TEE explanations are all possible combinations of these themes, following the structure above. Complete TEE explanations are shown in Appendix D.

Candidate TEE Explanation Performance We found that some TEE explanations performed significantly better than others for some of the comprehension questions. For example, for one of the questions about who is allowed to access their data (question 2 in Appendix B), participants were significantly more likely to answer the questions correctly if they received a *Non-Technical* TEE explanation (with $OR = 2.43, p < 0.05$ for the IoT scenario and $OR = 3.09, p < 0.01$ for the medical scenario, see Appendix E). Meanwhile, we saw no significant differences in the effect on trust.

We gave people the opportunity to ask questions about TEEs. The most common questions about TEEs were high-level, asking for more technical details (*"How exactly does a TEE work?"* (P55)). Other people wanted to know how they could be sure the TEE would work (*"... how [is] it guaranteed that it can't be accessed?"* (P87)). Yet others were curious if the technology was actually real (*"Is it a real thing? Or a hypothetical idea just for the study?"* (P127)). Interestingly, many people did not ask questions about TEEs at all, but instead used the space to share their thoughts about other topics. Some talked about our scenarios (*"... I was biased about this to begin with. I don't trust these devices"* (P62)). Others, about technology in general (*"You do understand that people don't trust technology?"* (P389)).

Future work Based on the questions asked by participants in our survey, we know that willingness to use technology and perception of data safety are more nuanced than can necessarily be communicated in a single Likert scale. We also learned that some people would benefit from additional technical details beyond our high-level explanations.

In a follow-up survey, we plan to answer some of the most frequently asked questions (**RQ4**) in the first study. A draft of the FAQ we will show to participants may be found in Appendix C. To evaluate the effectiveness of the FAQ, some participants will be shown the FAQ and encouraged to read its contents by imposing a 60-second wait on the survey. Others will be given the FAQ expandable menus, which they can choose to (or choose not to) expand and read. Another group will not receive the FAQ at all.

We will also refine our survey questions to better understand *which aspects* of our scenarios (such as the use of a TEE, the type of data being collected, and who is doing the collection) contribute to (or counteract) feelings of safety (**RQ5**). This new question will be evaluated on a 4-point Likert scale. In addition to asking if our participants have any lingering questions about TEEs, we will also ask them if there is *anything else* about the scenario that contributes to their belief that the data would be safe (or unsafe).

References

[1] How do I balance my sample within demographics? https://researcher-help.prolific.com/hc/en-gb/articles/360009221213-How-do-I-balance-my-sample-within-demographics. Accessed: 2024-05-17.

[2] Carolina Carreira, João F. Ferreira, Alexandra Mendes, and Nicolas Christin. Exploring usable security to improve the impact of formal verification: A research agenda. *Electronic Proceedings in Theoretical Computer Science*, 349:77–84, November 2021.

[3] Rachel Cummings, Gabriel Kaptchuk, and Elissa M Redmiles. "I need a better description": An investigation into user expectations for differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 2021.

[4] Steven Furnell, Rawan Esmael, Weining Yang, Ninghui Li, et al. Enhancing security behaviour by supporting the user. *Computers & Security*, 75:1–9, 2018.

[5] Hana Habib and Lorrie Faith Cranor. Evaluating the usability of privacy choice mechanisms. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, pages 273–289, 2022.

[6] Pratik Musale and Adam J Lee. Trust TEE?: Exploring the impact of trusted execution environments on smart home privacy norms. volume 3, 2023.

[7] Mohamed Sabt, Mohammed Achemlal, and Abdelmadjid Bouabdallah. Trusted execution environment: What it is, and what it is not. In *2015 IEEE Trustcom/BigDataSE/Ispa*, volume 1, 2015.

A Codebook

The codebook used to analyze the TEE explanations found in-the-wild is shown in Table 1.

B Survey Questions to Measure Understanding

To measure understanding we include the following True/False questions in our surveys. The correct answer is **bolded**. Text which differs between scenario will be shown in [brackets]. Questions 11 and 12, are for the follow-up survey.

Based on what you read above about TEEs, please tell us whether the following statements are true or false.

1. True/**False**: A member of the general public can access your data
2. True/**False**: [A hospital employee unrelated to the research team/Someone working at the company on an unrelated team] can access your data
3. True/**False**: If there were a bug in other software on the computer, outside of the TEE storing your data, then a hacker could use the bug to access your data
4. True/**False**: If a disgruntled [hospital employee unrelated to the research team/employee on an unrelated team] installed a malicious program on the computer storing your data, then they could access your data
5. True/**False**: Other [researchers/developers] working on different projects on the same computer can access your data
6. **True**/False: A member of the [research/development] team can access your data
7. **True**/False: If [a member of the research team/the light bulb/the voice assistant] makes a mistake collecting your data, then your data could be incorrect
8. **True**/False: [A member of the research team/Someone on the development team] could later use your data to [choose the location for a new fire station/train another AI diagnosis tool for a different medical condition/develop a smart vacuum]
9. True/**False**: The TEE ensures [the hospital being constructed will be closer to the patients who most need it/the new light bulbs will have features relevant to you/the diagnosis made by the AI tool will always be correct/your voice will always be recognized by the improved AI]
10. **True**/False: [A member of the research team/Someone on the development team] could steal your data and sell it on the dark web
11. **True**/False: When you unlock your Android phone with a PIN, the PIN is verified in a TEE
12. True/**False**: We cannot be sure that a TEE is configured correctly

C FAQ from Follow-Up Survey

In the follow-up survey, we answer some of the frequently asked questions from the first survey.

1. How do TEEs work? The details of how different TEEs work can vary. For example, Arm TrustZone is a feature of modern processors that splits computer resources between a "Normal World" and a "Trusted World" (a TEE). Software

Table 1: Codebook for "in-the-wild" explanations and respective frequency of each code. Each explanation could have up to 7 different codes.

Code	Description	Frequency
Reputation	Leverages pre-existing trust/reputation of recognizable companies	2
Verified	Application running in the TEE is verified	2
Attestation	Process to check that the software supporting the TEE is the code we expect	4
"Trust"	Explanation mentions the word "trust"	5
Unsubstantial	Generic/un-detailed description	8
Threat	TEE protects against untrusted OS/peripherals	8
Techniques	Describes particular TEE implementation (e.g., Intel SGX, Arm TrustZone)	10
Cryptography	Mentions the use of cryptography/cryptographic concepts	10
Technical	Explanation uses technical terminology (e.g., "confidentiality", "attestation")	11
Integrity	TEE prevents unauthorized modification	16
Prevents	TEE prevents some undesirable behavior	17
Secrecy	TEE prevents unauthorized access	21
Isolation	TEE ensures isolation from the rest of the system	23
Hardware	Mentions that a TEE is *hardware*-supported	23

running in each world has access to different regions of memory. Software running in the Normal World cannot access or modify data in the Trusted World. TrustZone is appropriate for protecting entire trusted applications while Intel SGX, on the other hand, works well with software that has both trusted and untrusted parts. SGX allows software to create one or more "enclaves" (TEEs). The data in the enclave can only be accessed while the trusted part of the software is running.

2. How do we know the TEE is working correctly? TEEs support hardware-based cryptographic functions that can be used to guarantee that both the TEE and all the code running in the TEE are configured properly. This process is called "attestation". Researchers are also working on new ways to ensure that software running in a TEE works as expected.

3. How are TEEs used in real life? TEES are used in computers, smart phones, and other devices. For example, authentication in modern Android phones is typically handled by code (called "Gatekeeper") residing in a TEE based on ARM TrustZone. For example, when you enter a PIN or scan your fingerprint to unlock your phone, it is sent to GateKeeper in the secure zone of the CPU to verify. The response from GateKeeper is encrypted with a secret, hardware-backed key that is never shared outside the TEE.

D Candidate TEE Explanations

Candidate TEE explanations are composed of 2-3 sentences, where each sentence has a different theme. We evaluate every combination of the 2-3 sentences in our surveys. Each theme is shown below in *italics*, followed by the corresponding sentence from our evaluation.

Sentence 1: Introducing TEEs
Hardware: A Trusted Execution Environment (TEE) is a technique for running programs and interacting with data securely using a protected area of the physical computer.
Trust: A Trusted Execution Environment (TEE) is a technique for running programs and interacting with data securely, even if the rest of the computer is not trustworthy.
Unsubstantial: A Trusted Execution Environment (TEE) is a technique for running programs and interacting with data securely.

Sentence 2: Isolation, confidentiality, and integrity
Technical: A program running in a TEE is isolated from the rest of the computer to protect the confidentiality and integrity of the program and data.
Non-Technical: A program running in a TEE is isolated from the rest of the computer to allow only authorized people to view or change the program and data.

Sentence 3: (Optional) threat prevented by TEE
Prevents: The TEE protects the program and data even when other software on the computer is behaving maliciously.
No Prevents: (No third sentence)

E Regressions and other results

Comprehension scores (overall and splity by scenario) may be found in Table 2. A similar table of comprehension scores split by TEE explanation may be found in Table 3. Finally, the regression table for the comprehension questions is shown in Table 4.

Q#	T/F	Overall	Medical Complex	Medical Simple	Smart Home Complex	Smart Home Simple
Features of TEEs						
Q1	F	96.7%	96.2%	97.0%	96.6%	97.0%
Q2	F	91.9%	92.4%	93.1%	90.6%	91.5%
Q3	F	79.0%	80.9%	76.8%	78.1%	80.1%
Q4	F	80.9%	83.1%	81.5%	80.7%	78.4%
Q5	F	84.2%	83.5%	84.1%	87.1%	82.2%
Limitations of TEEs						
Q6	T	87.2%	91.1%	87.6%	83.3%	86.9%
Q7	T	82.0%	86.9%	80.3%	82.0%	78.8%
Q8	T	69.8%	87.3%	63.1%	66.1%	62.7%
Q9	F	61.1%	82.2%	57.1%	54.1%	50.8%
Q10	T	57.2%	59.3%	55.4%	61.4%	53.0%

Table 2: Overall scores for each comprehension question highlighting features and limitations of TEEs and the correct answers. Results are split by scenario.

			First sentence			Second sentence		Third sentence	
	T/F	Overall	Hardware	Trust	Unsubstantial	Technical	Non-technical	Prevents	No Prevents
Features of TEEs									
Q1	F	96.7%	97.8%	94.5%	97.8%	96.6%	96.8%	95.7%	97.7%
Q2	F	91.9%	94.0%	87.7%	93.9%	88.7%	95.1%	91.5%	92.3%
Q3	F	79.0%	78.6%	81.8%	76.6%	78.4%	79.6%	82.5%	75.5%
Q4	F	80.9%	80.5%	80.8%	81.4%	82.7%	79.1%	87.8%	74.0%
Q5	F	84.2%	84.6%	82.5%	85.6%	82.9%	85.5%	85.9%	82.6%
Limitations of TEEs									
Q6	T	87.2%	88.1%	87.3%	86.2%	86.1%	88.3%	87.2%	87.2%
Q7	T	82.0%	79.6%	82.8%	83.7%	84.0%	80.0%	81.6%	82.3%
Q8	T	69.8%	67.9%	73.7%	67.9%	69.2%	70.4%	71.2%	68.5%
Q9	F	61.1%	56.9%	60.1%	66.3%	62.8%	59.4%	60.5%	61.7%
Q10	T	57.2%	57.5%	58.1%	56.1%	53.0%	61.5%	53.8%	60.6%

Table 3: Overall scores for each comprehension question in Survey 1, highlighting features and limitations of TEEs and the correct answers. Results are split by TEE explanation.

Simple Medical Research Scenario										
Variable	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Explanation sentence 1 [Baseline = Unsubstantial]										
Hardware	0.42	−0.96	0.28	−0.01	−0.85	0.03	−0.76	−0.09	−0.90**	0.25
Trust	0.38	−0.62	0.68	−0.01	−0.58	0.41	−0.35	0.41	−0.24	−0.07
Explanation sentence 2 [Baseline = Technical]										
Non-technical	0.16	2.14**	−0.07	0.17	0.33	0.49	−0.12	0.14	0.08	0.20
Explanation sentence 3 [Baseline = No Prevents]										
Prevents	−0.38	0.10	0.64*	1.30***	0.04	−0.11	−0.30	−0.01	0.13	−0.64*
Medical experience	−1.43	−0.08	0.09	0.75	0.29	−0.18	0.02	0.02	0.27	0.34
CS experience	−0.31	0.49	−0.03	0.19	0.35	0.16	0.08	0.53	0.37	−0.07
Complex Medical Research Scenario										
Variable	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Explanation sentence 1 [Baseline = Unsubstantial]										
Hardware	0.03	1.60	0.07	0.11	0.11	1.04	0.62	0.41	−0.13	−0.31
Trust	−1.77	−0.78	−0.15	−0.18	0.15	0.36	0.49	0.18	−0.18	0.04
Explanation sentence 2 [Baseline = Technical]										
Non-technical	−0.24	0.47	0.10	−0.47	0.43	−0.12	−0.47	−0.49	−0.55	0.71*
Explanation sentence 3 [Baseline = No Prevents]										
Prevents	−0.56	−0.93	0.53	1.04**	0.24	−0.04	0.56	0.51	−0.61	0.01
Medical experience	0.32	0.79	0.21	0.01	−0.23	0.27	−0.05	1.01	1.13	1.28***
CS experience	−1.45*	−0.82	−0.14	−0.78*	−0.95*	−0.07	0.24	0.40	−0.23	0.26
Simple Smart Home Device Scenario										
Variable	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Explanation sentence 1 [Baseline = Unsubstantial]										
Hardware	−0.04	−0.65	−0.28	−0.35	0.02	0.19	−0.10	−0.11	−0.08	0.32
Trust	−0.48	−0.46	0.33	0.19	−0.29	−0.07	−0.34	−0.01	−0.13	0.19
Explanation sentence 2 [Baseline = Technical]										
Non-techncial	−0.18	0.94	0.23	−0.08	0.04	0.28	−0.14	0.11	−0.25	0.24
Explanation sentence 3 [Baseline = No Prevents]										
Prevents	−0.99	−0.01	0.07	0.51	0.16	0.03	−0.53	0.03	0.11	0.01
IOT experience	−15.97	−1.46	−1.26*	−1.02	−0.46	−0.57	−0.51	−0.32	−0.79*	−0.03
CS experience	−1.76*	0.41	−0.36	−0.52	−0.16	−1.00*	−0.05	−0.68*	0.66	−0.43
Complex Smart Home Device Scenario										
Variable	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Explanation sentence 1 [Baseline = Unsubstantial]										
Hardware	−0.81	−0.10	0.31	0.39	0.29	−0.35	−0.41	−0.11	−0.49	−0.42
Trust	−1.72	−1.06	0.35	−0.21	−0.22	−0.22	0.16	0.40	−0.70*	0.02
Explanation sentence 2 [Baseline = Technical]										
Non-technical	0.13	0.71	0.14	−0.36	−0.05	0.02	−0.61	0.21	0.11	0.35
Explanation sentence 3 [Baseline = No Prevents]										
Prevents	−0.60	0.73	−0.39	0.99**	0.80	0.15	0.26	0.11	−0.25	−0.40
IoT experience	−0.60	0.73	−0.39	0.23	0.18	−0.65	0.06	−0.36	0.01	−0.49
CS experience	−0.79	−0.46	0.16	0.23	−0.39	0.07	0.33	0.24	0.42	0.29

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 4: Regression table for True/False comprehension questions. There is one logistic regression model for each question in each scenario (40 models total). The numbers in this table are the log-odds coefficients for each predictor, with the baseline explanations used in each model noted in *italics*. Statistical significance is noted with asterisks and shaded cells: blue for positive coefficients and orange for negative coefficients.