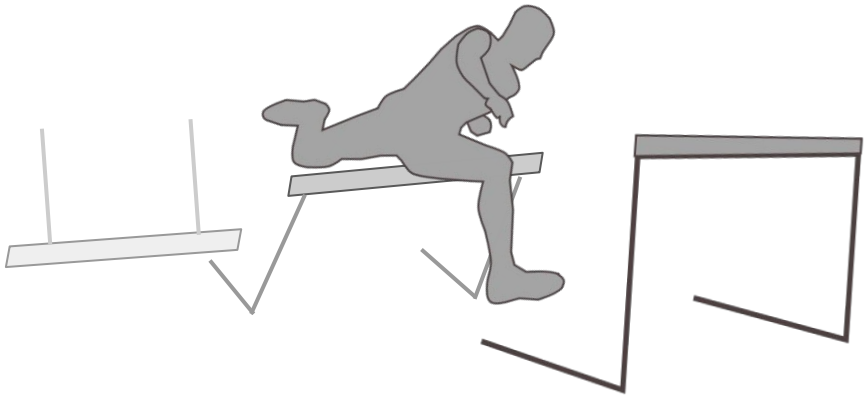
Evaluating In-Workflow Messages for Improving Mental Models of End-to-End Encryption

Omer Akgul • Wei Bai • Shruti Das • Michelle L. Mazurek

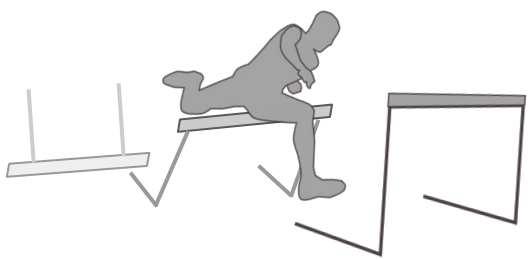
Adoption of E2EE By **General Users**?

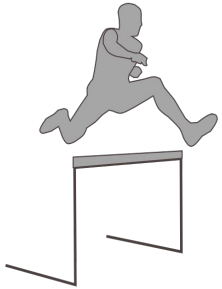
Many hurdles impede adoption!



Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

Many Hurdles Impede Adoption

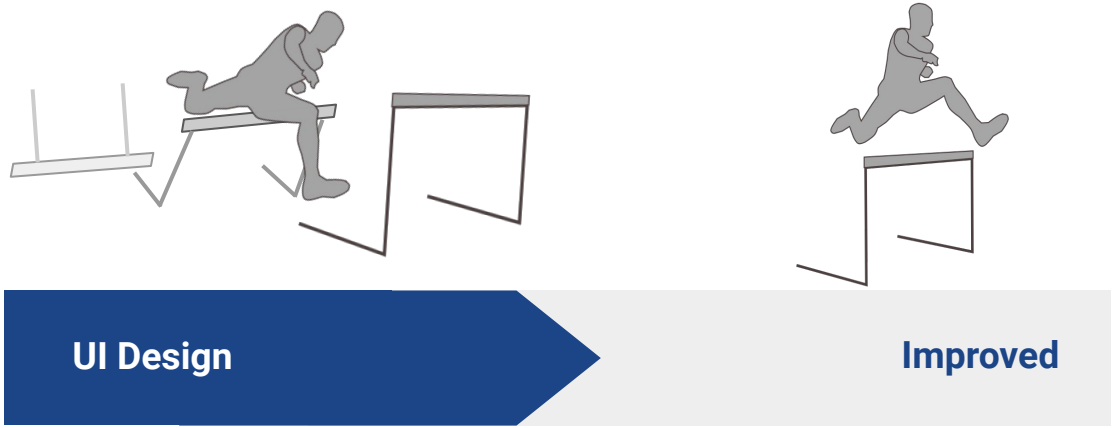




UI Design

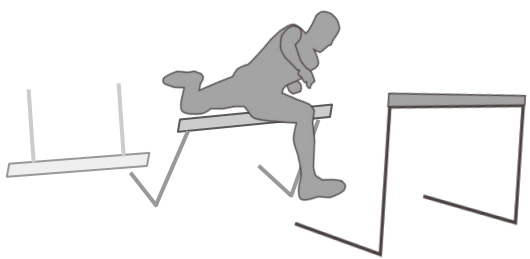
Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

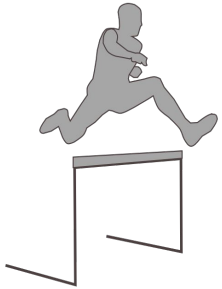
Many Hurdles Impede Adoption



Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

Many Hurdles Impede Adoption



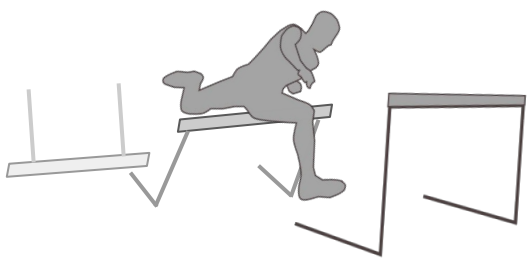


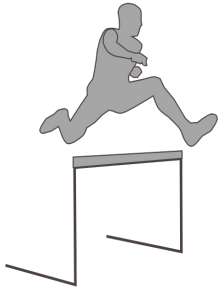
UI Design → **Improved**

Key Management

Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

Many Hurdles Impede Adoption

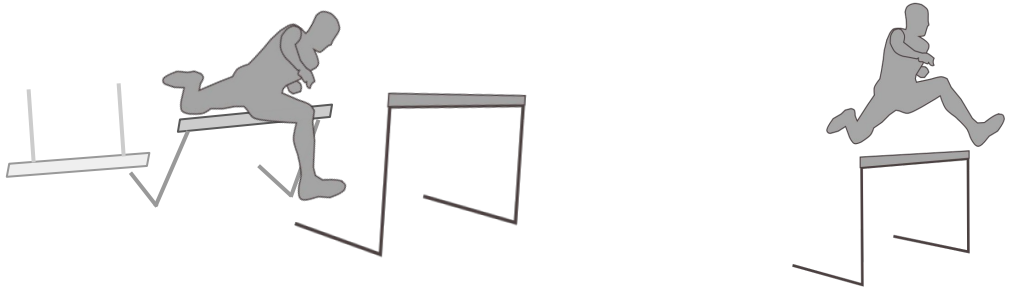




UI Design	Improved
Key Management	Key-Directory Based Model

Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

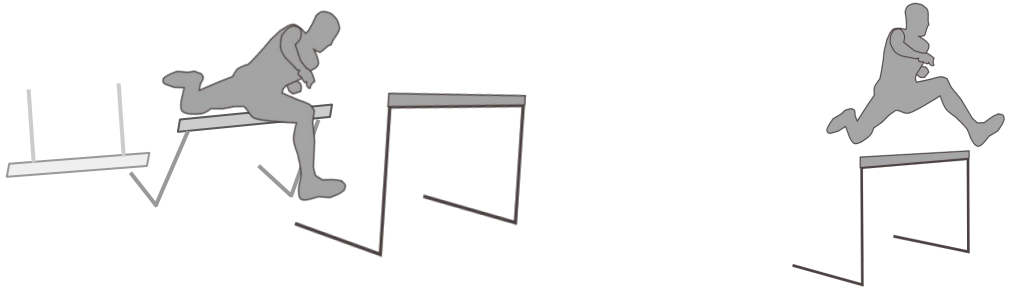
Many Hurdles Impede Adoption



UI Design	Improved
Key Management	Key-Directory Based Model
Social Norms	

Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

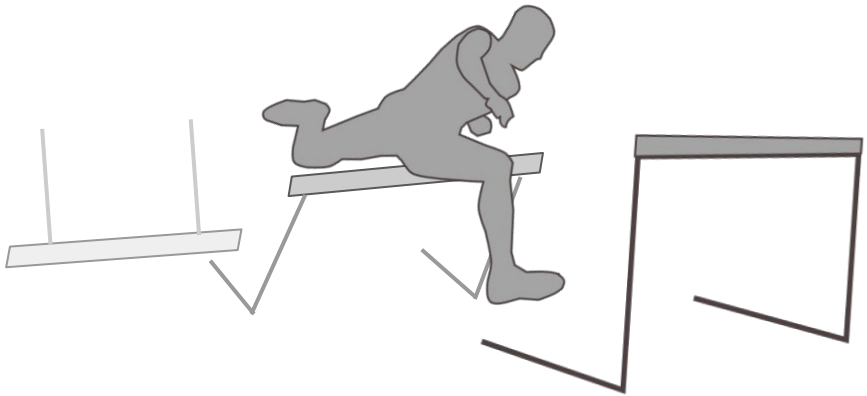
Many Hurdles Impede Adoption



UI Design	Improved
Key Management	Key-Directory Based Model
Social Norms	Large Deployment

Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

But... Mental Models still a problem!



Icons from https://freesvg.org/vector-silhouette-of-an-athlete with modification

Why do (incorrect) mental models matter?

People perceive E2EE incorrectly in both directions [1-2]:

- Encryption protects from anything
- Encryption can be trivially broken by anyone who works in IT

[1] Abu-Salma et al. Obstacles to the adoption of secure communication tools. In IEEE Security & Privacy, 2017
[2] Wu et al. When is a Tree Really a Truck? Exploring Mental Models of Encryption. In SOUPS 2018

Icon from https://pixabay.com/illustrations/question-question-mark-response-1015308/

Why do (incorrect) mental models matter?

People perceive E2EE incorrectly in both directions [1-2]:

- Encryption protects from anything
- Encryption can be trivially broken by anyone who works in IT

Difficult for users to make thoughtful decisions:

- “SMS is the most secure messaging service.” [1]

[1] Abu-Salma et al. Obstacles to the adoption of secure communication tools. In IEEE Security & Privacy, 2017
[2] Wu et al. When is a Tree Really a Truck? Exploring Mental Models of Encryption. In SOUPS 2018

Icon from https://pixabay.com/illustrations/question-question-mark-response-1015308/

Why do (incorrect) mental models matter?

People perceive E2EE incorrectly in both directions [1-2]:

- Encryption protects from anything
- Encryption can be trivially broken by anyone who works in IT

Difficult for users to make thoughtful decisions:

- "SMS is the most secure messaging service." [1]

Because they inhibit

Confident, Proactive, and Correct usage

[1] Abu-Salma et al. Obstacles to the adoption of secure communication tools. In IEEE Security & Privacy, 2017
[2] Wu et al. When is a Tree Really a Truck? Exploring Mental Models of Encryption. In SOUPS 2018

Icon from https://pixabay.com/illustrations/question-question-mark-response-1015308/

Improve mental models **Naturally**

Goal: **Help people grok basic understanding and threats**

- **Enough** to make judgments about how to communicate
- **Without** turning everyone into crypto experts
- **Without** requiring people to sign up for training modules

Solution: **Place educational messages in a messaging app, where people see them.**

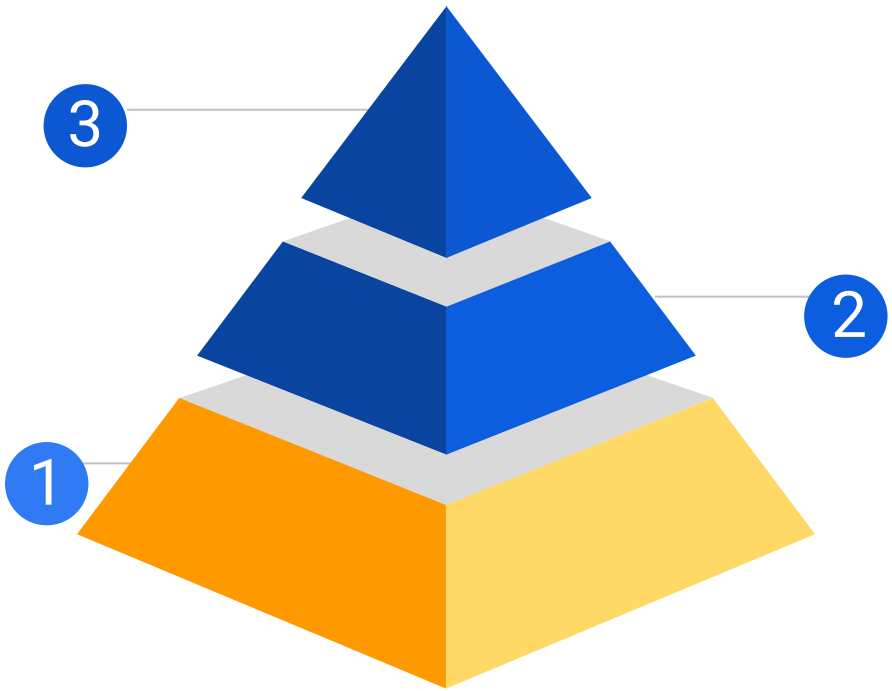
Multi-Stage Efforts: From Lab to Field

Field(ish) Study

- Fit messages to an app
- Daily use for 3 weeks

Lab Study

- In-depth tutorial
- What's important, difficult?



Online Survey

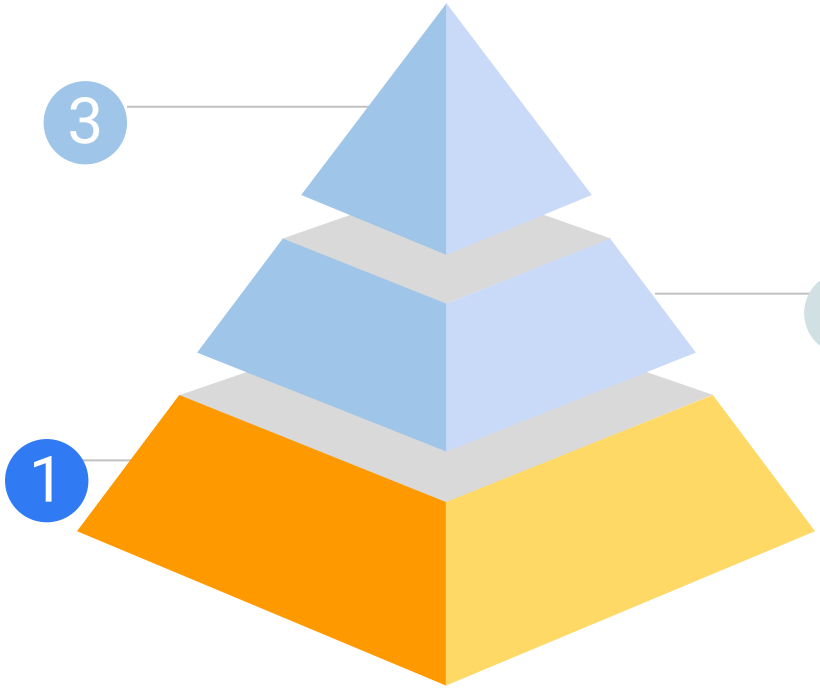
- Test different messages varying in length and contents

Bai et al. Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study. In EuroUSEC, 2020

Multi-Stage Efforts: From Lab to Field

Field(ish) Study

- Fit messages to an app
- Daily use for 3 weeks



Online Survey

- Test different messages varying in length and contents

Lab Study

- In-depth tutorial

Bai et al. Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study. In EuroUSEC, 2020

Study 1 - Takeaways

- Confidentiality: Most significant

- Explaining risks clearly is useful
  - Comparing E2EE vs Non-E2EE
  - Weakness

- Some pieces may not worth mentioning
  - Integrity & authenticity
  - How E2EE works

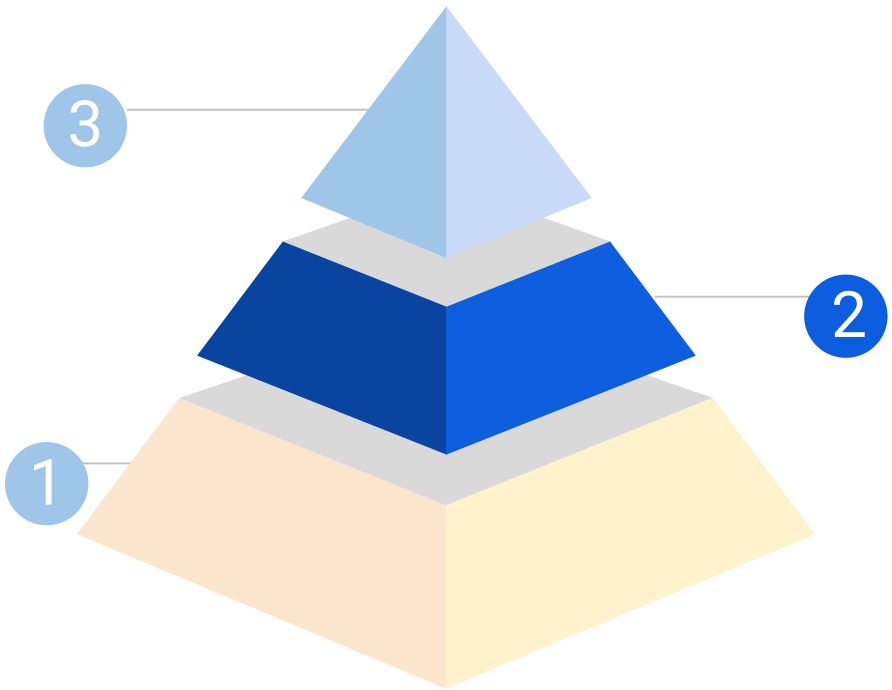
Multi-Stage Efforts: From Lab to Field

Field(ish) Study

- Fit messages to an app
- Daily use for 3 weeks

Lab Study

- In-depth tutorial



Online Survey

- Test different messages varying in length and contents

Bai et al. Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study. In EuroUSEC, 2020

Testing educational messages

- Can we **shift** user mental models with short messages in text?
  - In isolation
- How much is lost in **short**, **medium** vs. **long** messages?
  - Appropriate for various UIs.
- Which short, medium messages are most effective (for what)?
- Don’t want to **oversell** security

Study 2: Setup

- Online study via a crowdsourcing platform (Prolific, n=461)
- 1 Long, 5 short, 2 medium, 1 control message
  - Hypothetical app called TextLight (to remove brand bias)
- One message per participant.

Message types

Short	

Message types

Short	(1) Nobody but you and recipient

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other
Medium	

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other
Medium	Two messages with various combinations of short messages.

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other
Medium	Two messages with various combinations of short messages.
Long	

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other
Medium	Two messages with various combinations of short messages.
Long	All key points, extra emphasis

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other
Medium	Two messages with various combinations of short messages.
Long	All key points, extra emphasis
Control	

Message types

Short	(1) Nobody but you and recipient (2) Metadata risks (3) Endpoint risks (4) Lock/key metaphor (5) E2EE vs. other
Medium	Two messages with various combinations of short messages.
Long	All key points, extra emphasis
Control	Describes non-security/privacy features

UNIVERSITY OF MARYLAND

Based on your understanding of end-to-end encryption, please indicate whether you agree or disagree that **hackers who have compromised the TextLight servers** have the following abilities, regardless of their motivation to do so.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Can see that you have sent a message on TextLight, regardless of knowing the content of the message.	○	○	●	○	○
Can see what is in the	○	○	●	○	○

Pre survey → Read Message → Exit survey

“Messages in TextLight are end-to-end encrypted. Before a message ever leaves your device, it’s secured with a lock, and only you and your recipients have the keys to open the message and read it.”

Pre survey | Read Message | Exit survey

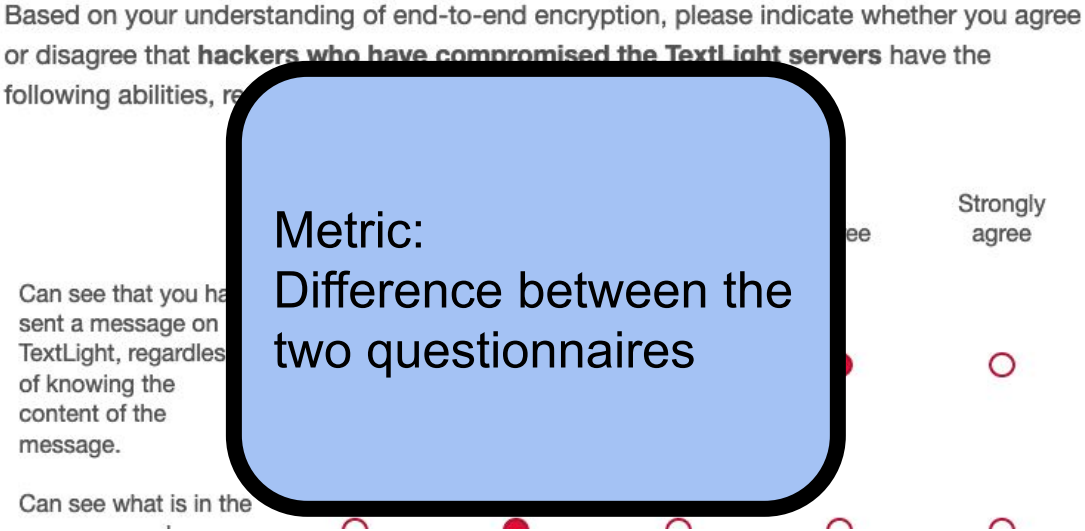
UNIVERSITY OF MARYLAND

Based on your understanding of end-to-end encryption, please indicate whether you agree or disagree that **hackers who have compromised the TextLight servers** have the following abilities, regardless of their motivation to do so.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Can see that you have sent a message on TextLight, regardless of knowing the content of the message.	○	○	○	●	○
Can see what is in the	○	●	○	○	○

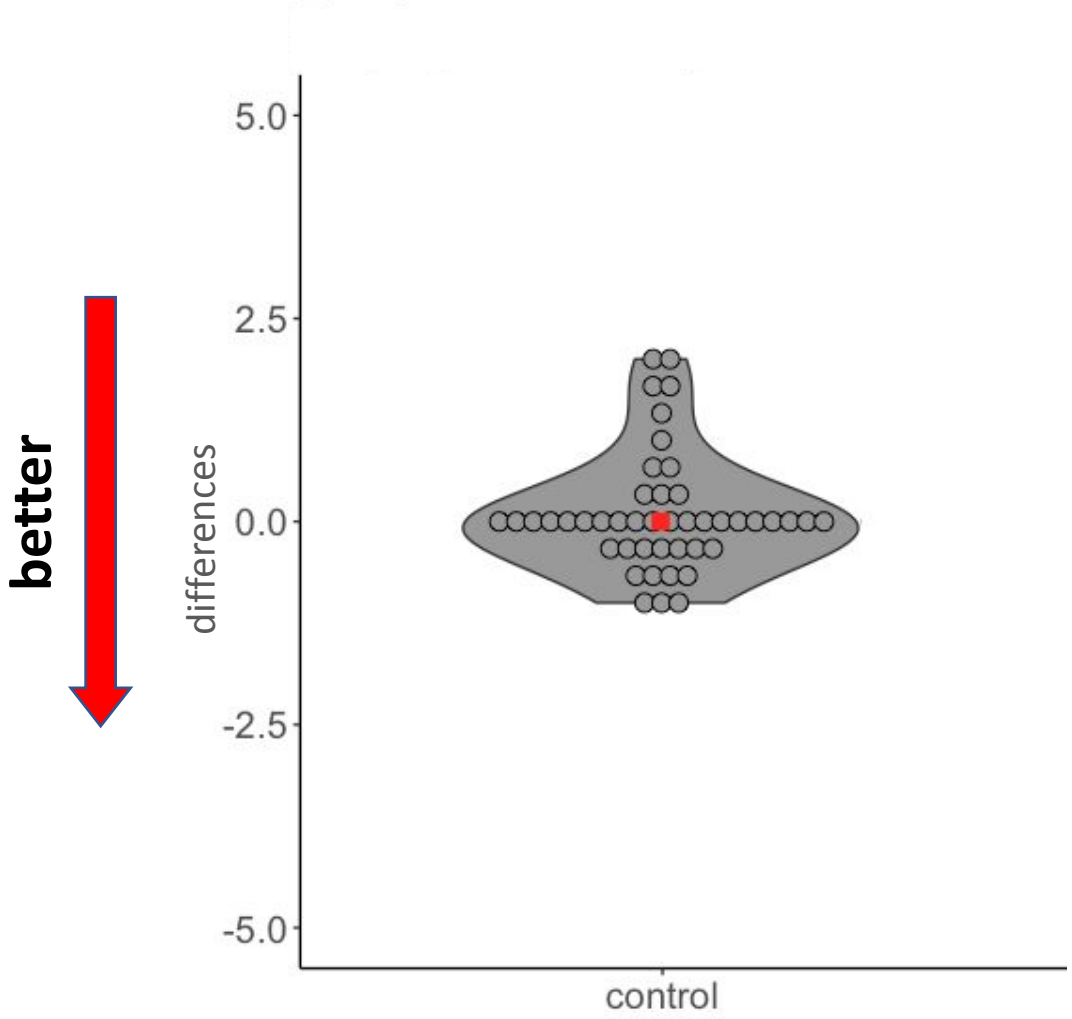
Pre survey → Read Message → Exit survey

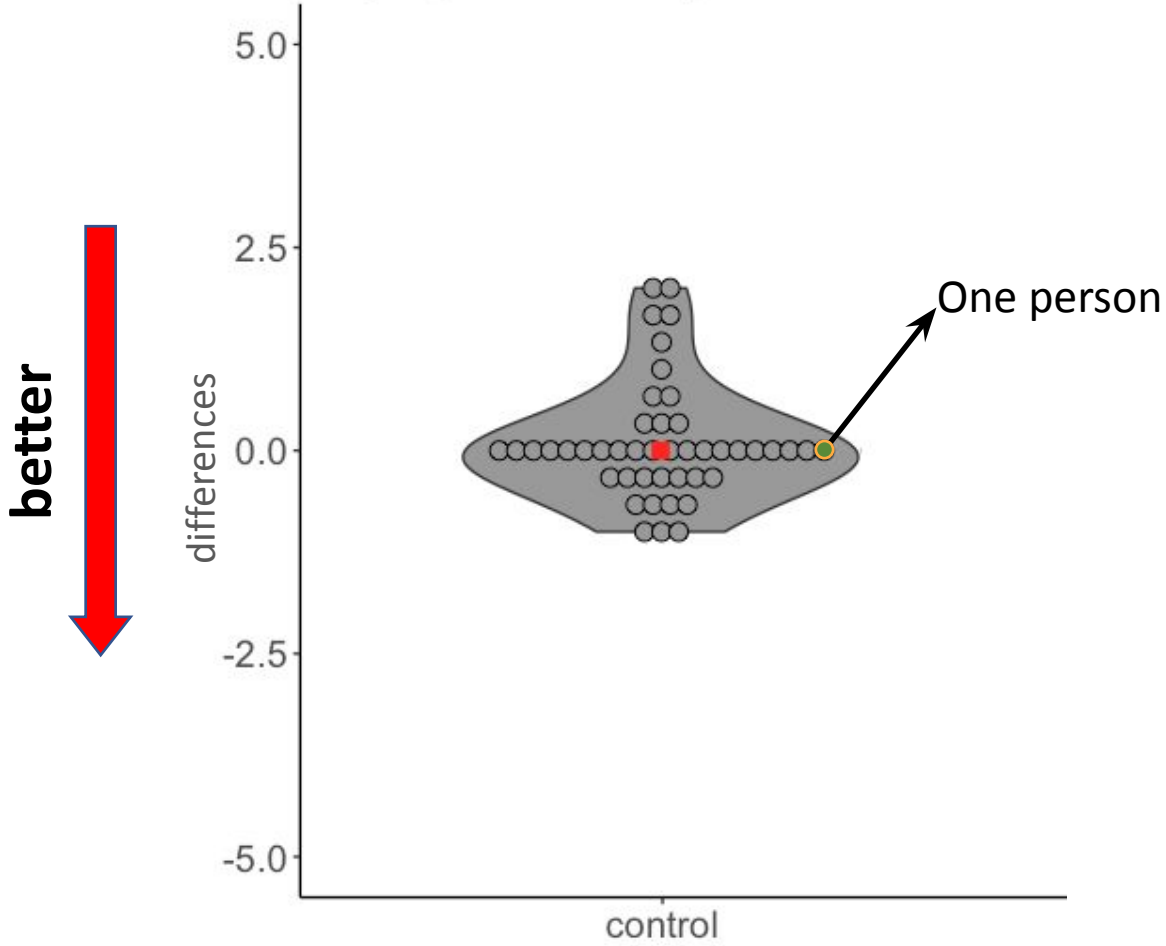
UNIVERSITY OF MARYLAND

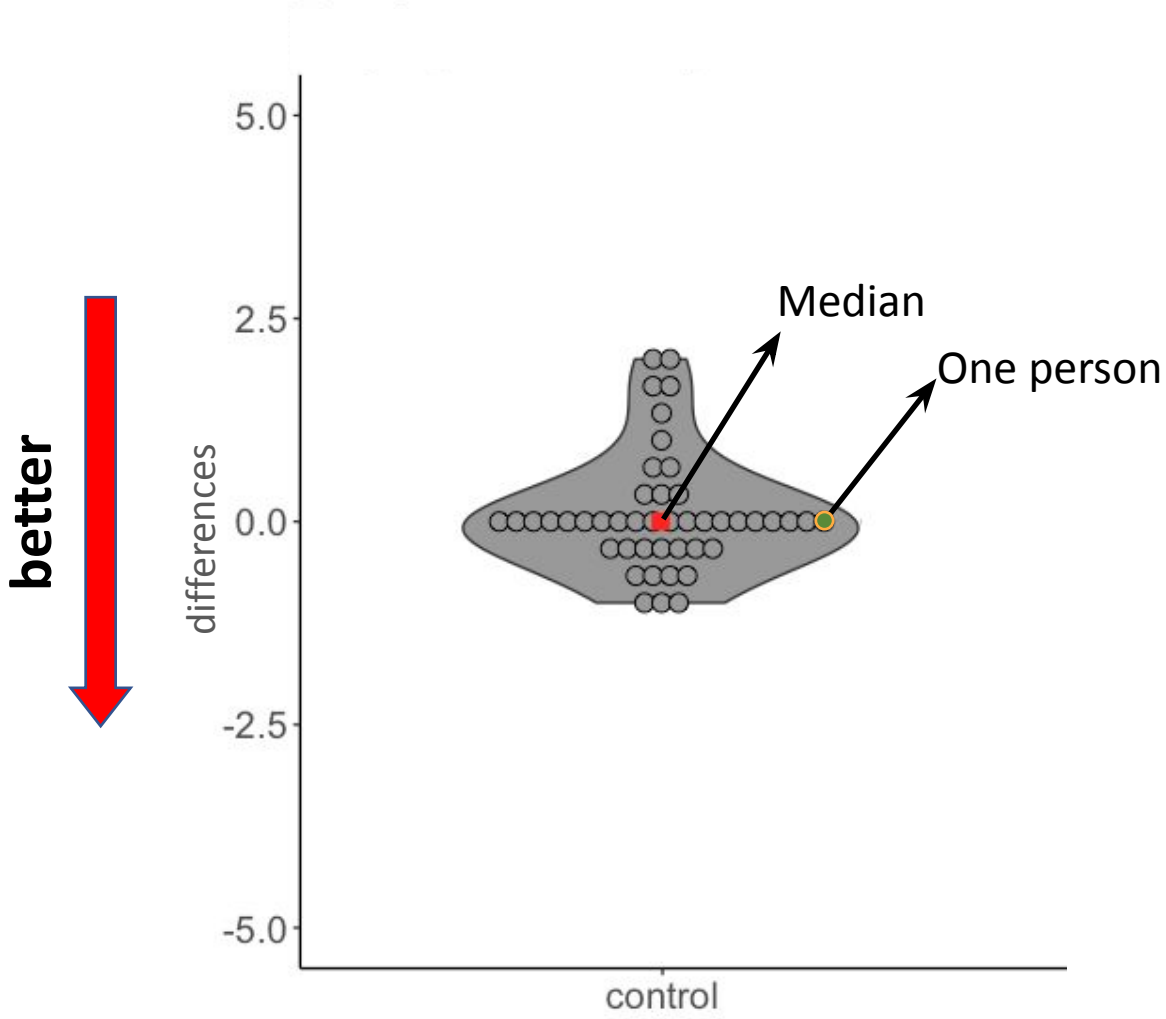


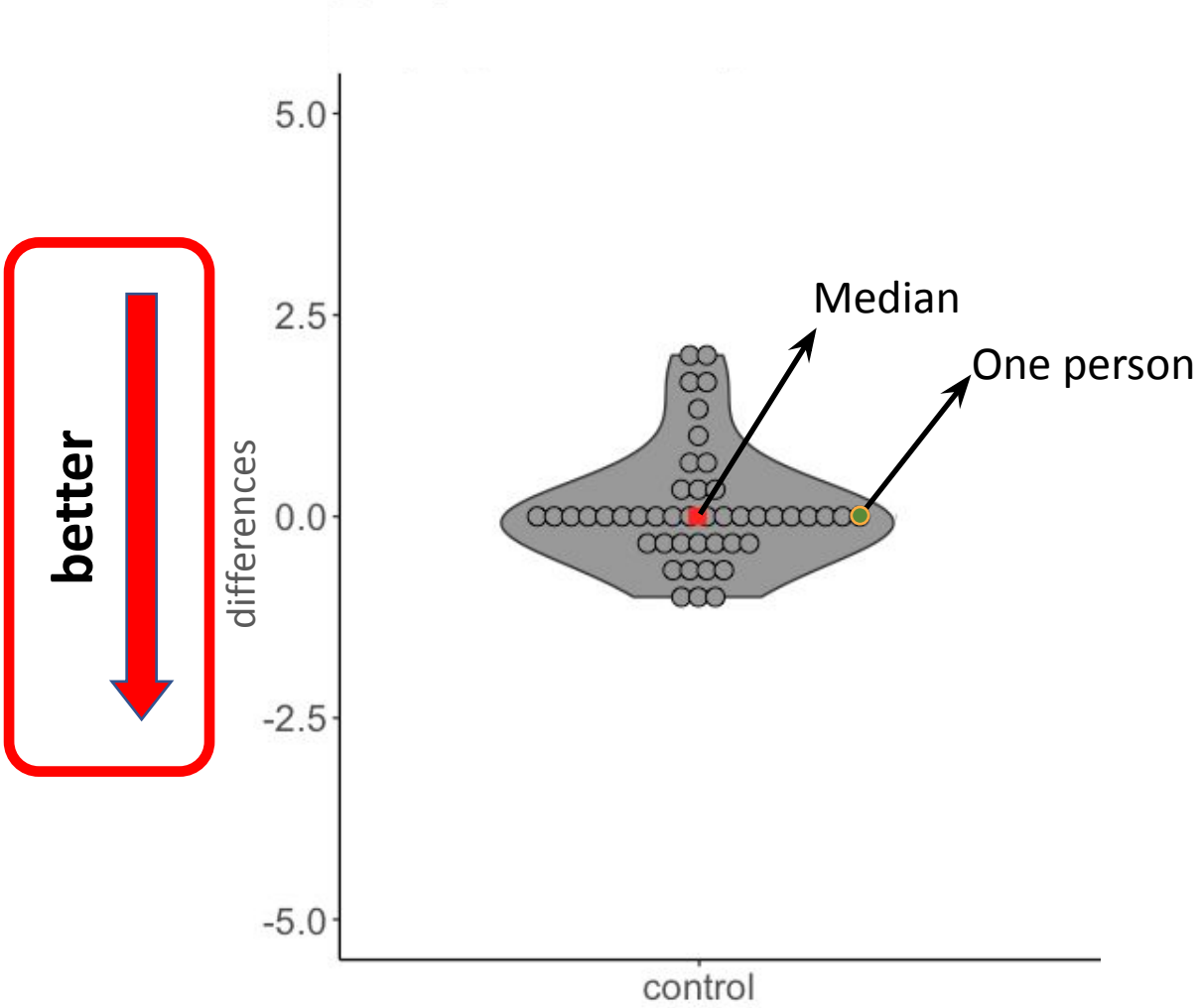
Metric:
Difference between the two questionnaires

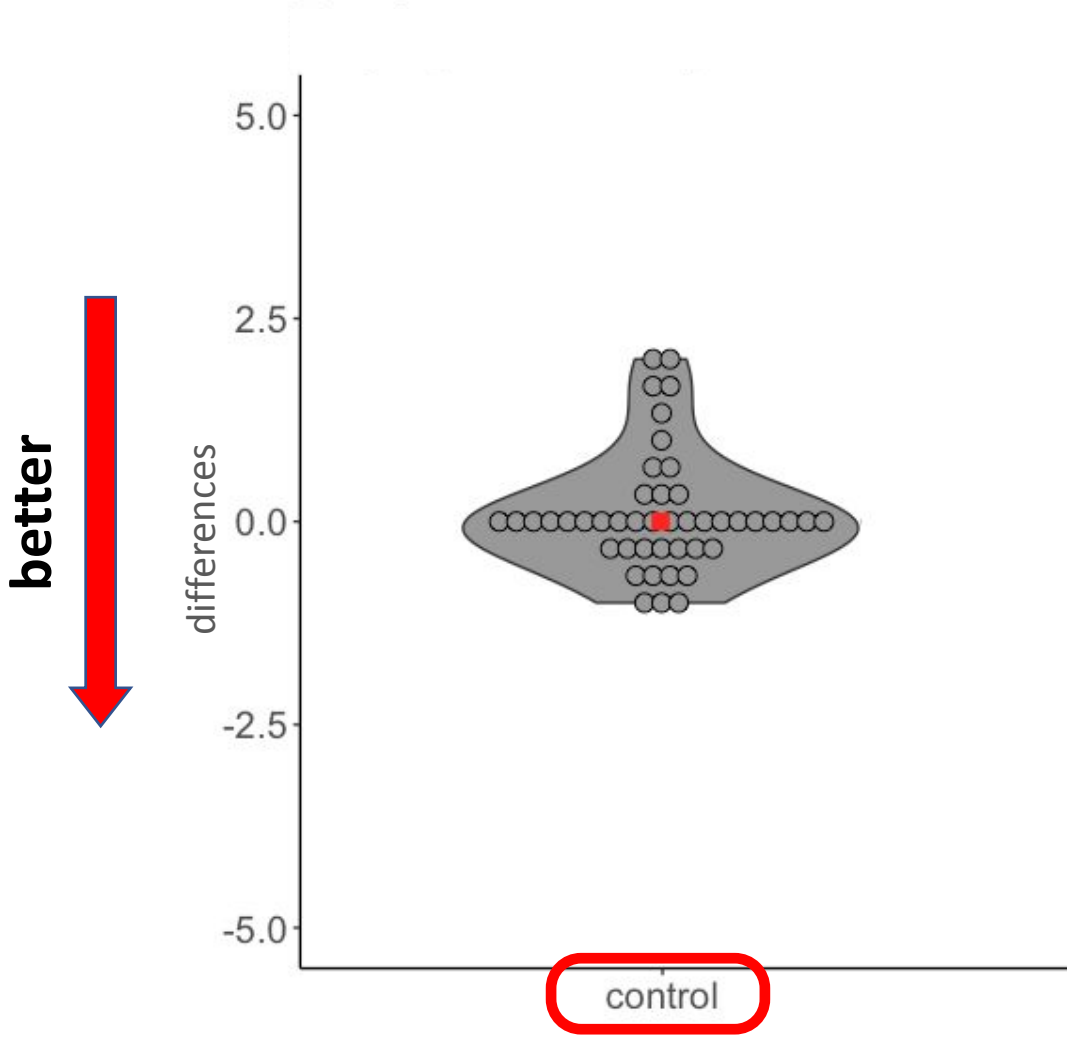
Pre survey → Read Message → Exit survey


better
differences
5.0
2.5
0.0
-2.5
-5.0
control


better
differences
5.0
2.5
0.0
-2.5
-5.0
control
One person

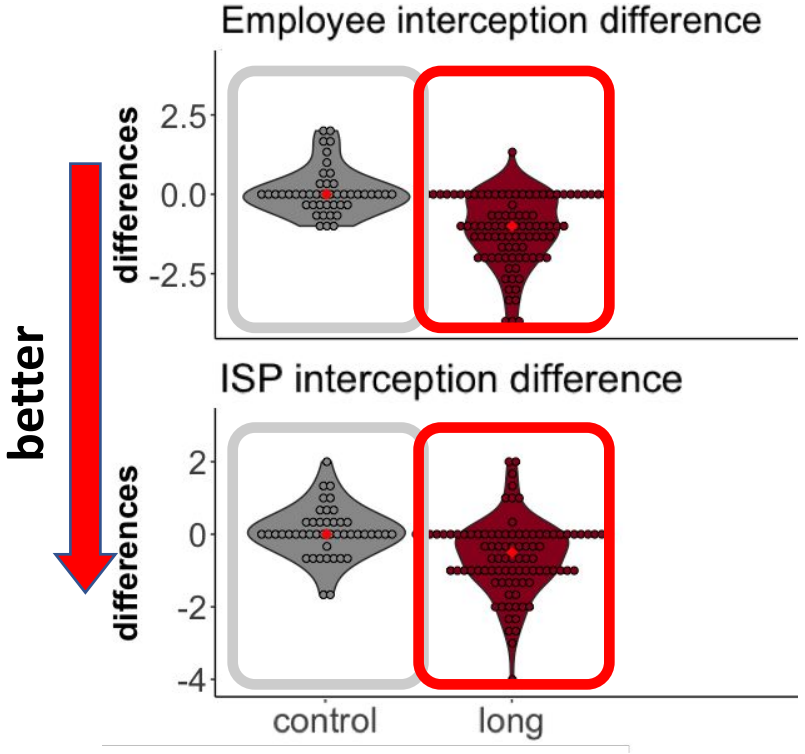

better
differences
5.0
2.5
0.0
-2.5
-5.0
control
Median
One person


better
differences
5.0
2.5
0.0
-2.5
-5.0
control
Median
One person


better
differences
5.0
2.5
0.0
-2.5
-5.0
control

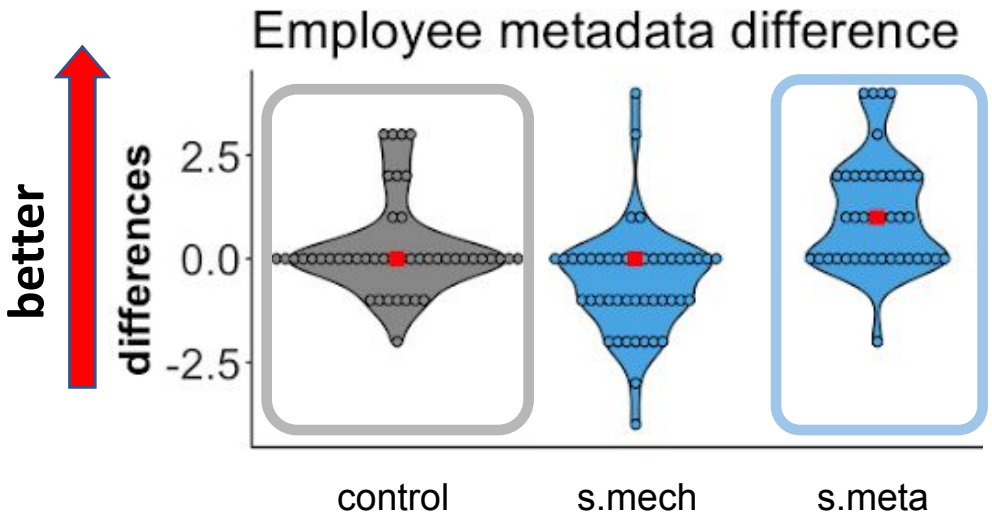
Long messages work!

- Long message is generally better than control
  - Our best effort



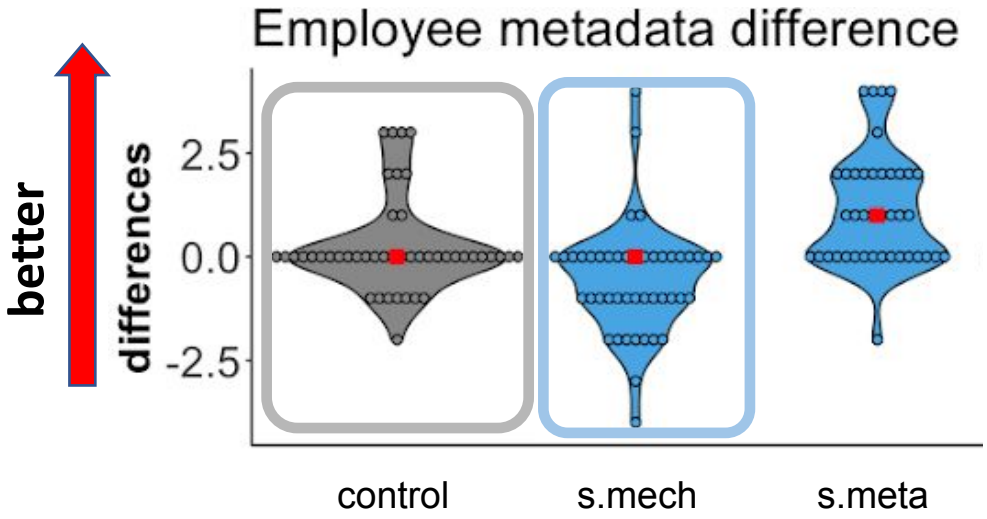
Shorts? Also work!

- When message is topical, mostly better than all messages



Shorts? Also work!

- When message is topical, mostly better than all messages
- But, some additional risk of overselling!



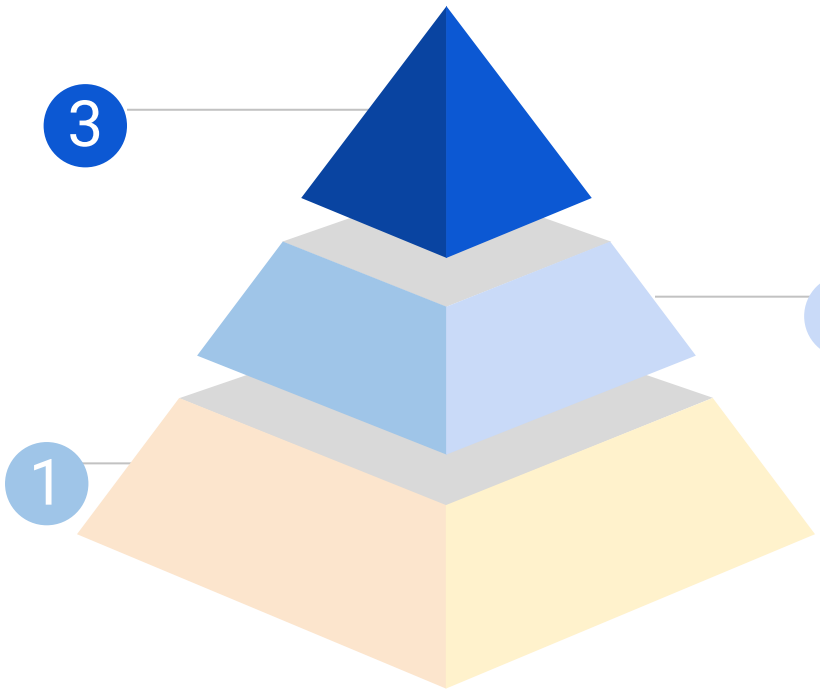
Study 2: Takeaways

- The messages work! (in a controlled environment)
- Short messages work surprisingly well
  - Chance of overselling, need all for a complete mental model

Multi-Stage Efforts: From Lab to Field

Field(ish) Study

- Fit messages to an app
- Daily use for 3 weeks



Online Survey

- Test different messages varying in length and contents

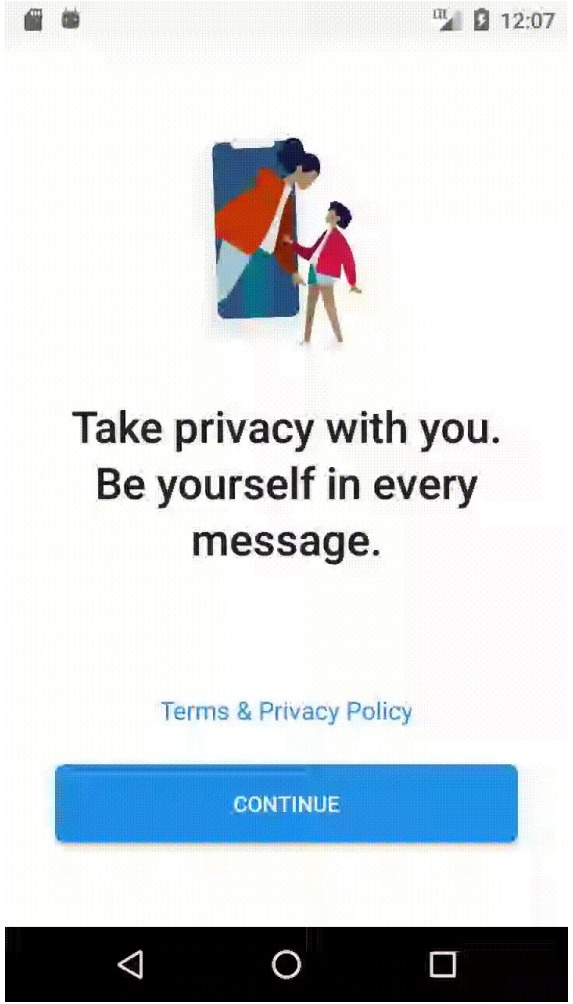
Lab Study

- In-depth tutorial

Bai et al. Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study. In EuroUSEC, 2020

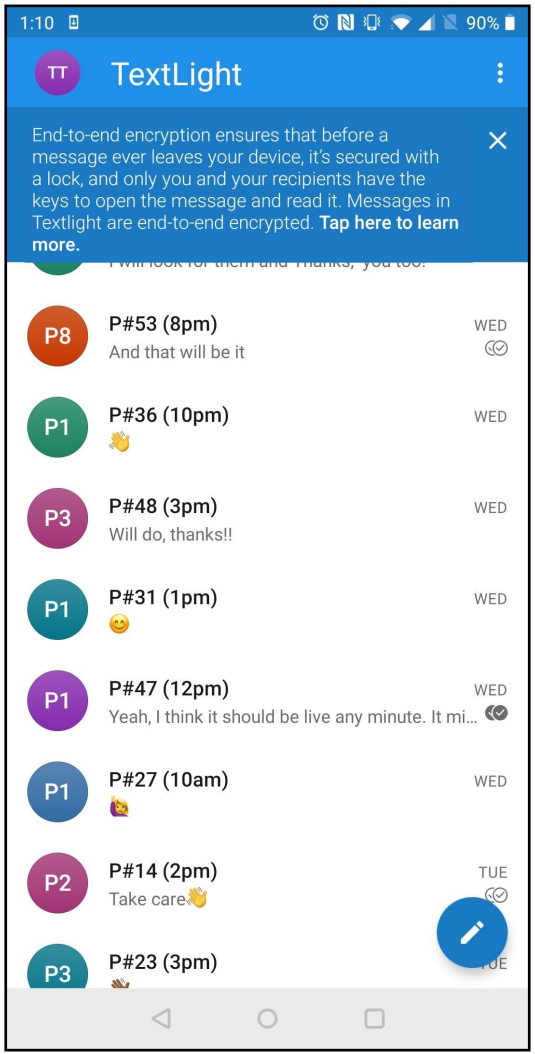
Feeds Into Study 3

- How well would messages from study 2 work in the real world?
  - (integrated in an app)
- Why does it or why doesn't it work?
  - How can we improve it further?



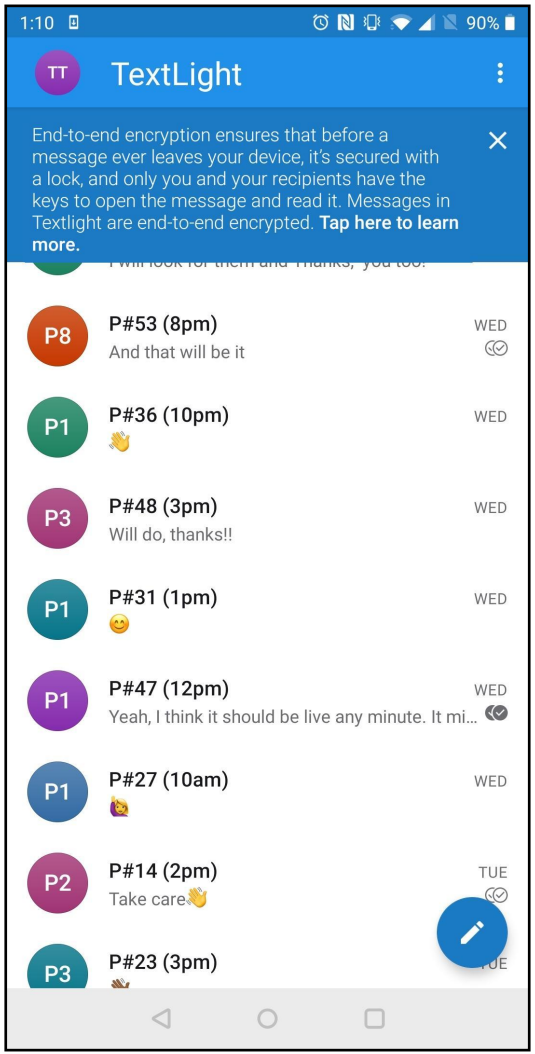
Study 3 Setup

- Incorporate successful messages from online study into an app (experimental)



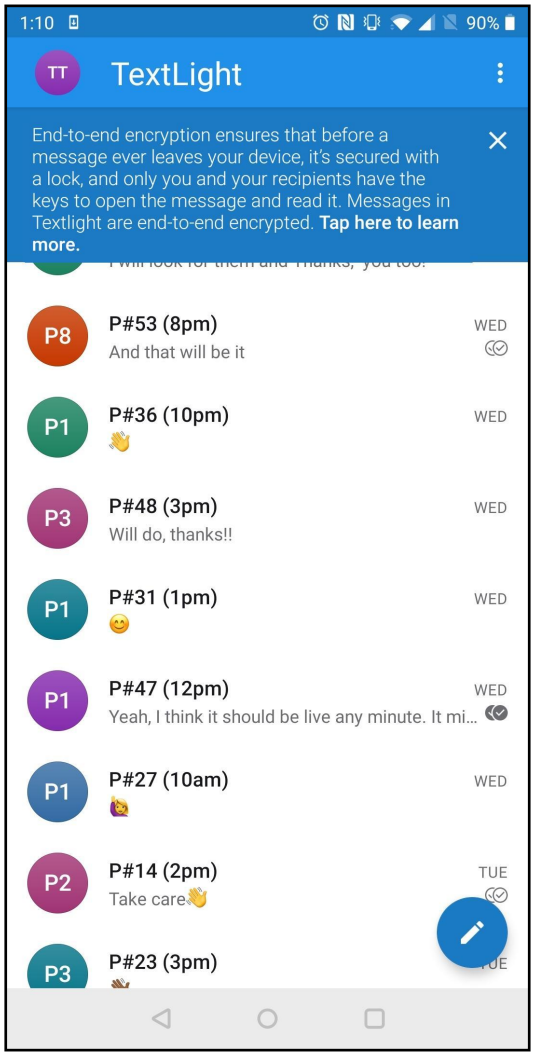
Study 3 Setup

- Incorporate successful messages from online study into an app (experimental)
  - Re-brand Signal to TextLight



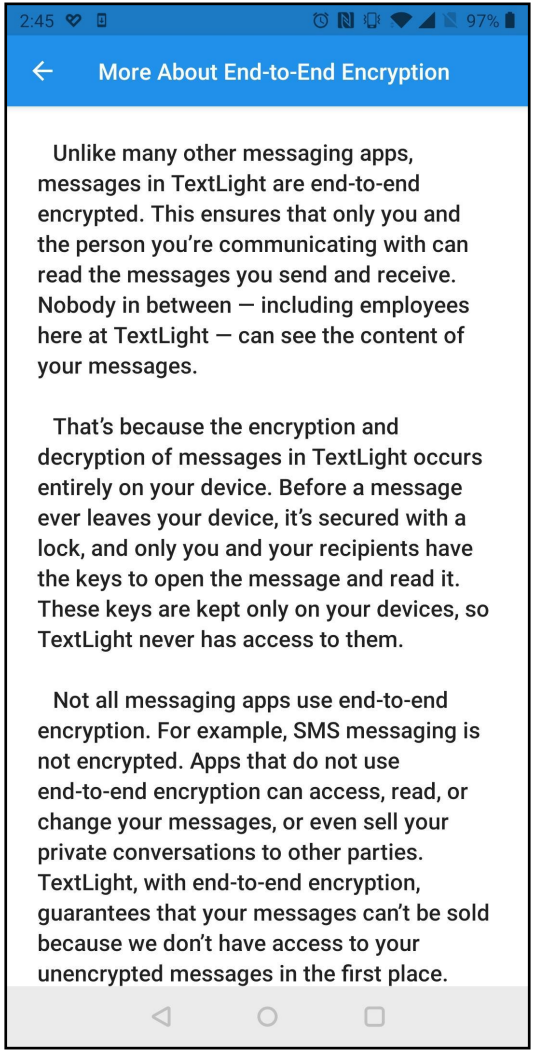
Study 3 Setup

- Incorporate successful messages from online study into an app (experimental)
  - Re-brand Signal to TextLight
  - Show short messages



Study 3 Setup

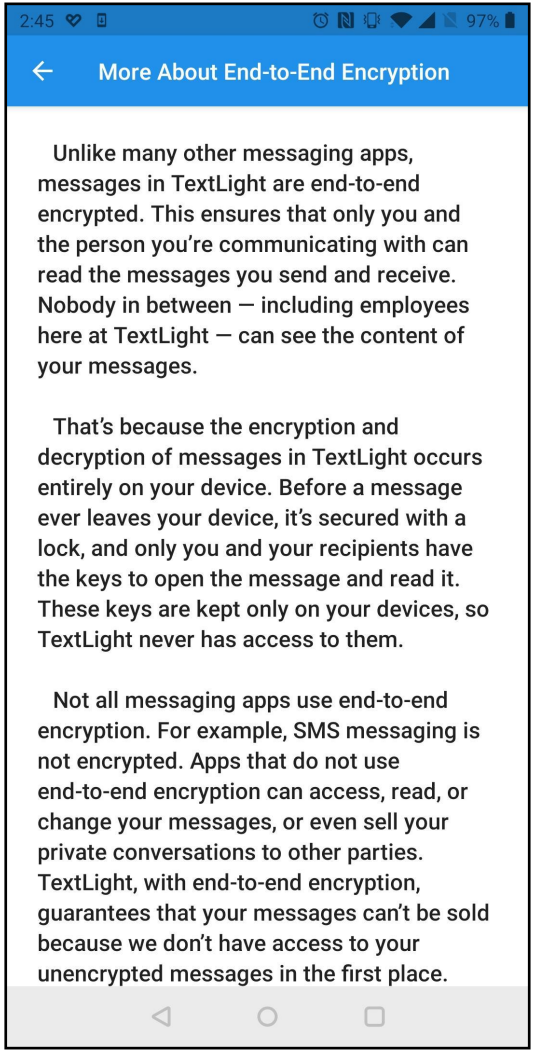
- Incorporate successful messages from online study into an app (experimental)
  - Re-brand Signal to TextLight
  - Show short messages
  - Clickable to open long message



Study 3 Setup

- Incorporate successful messages from online study into an app (experimental)
  - Re-brand Signal to TextLight
  - Show short messages
  - Clickable to open long message

- Control version with no messages

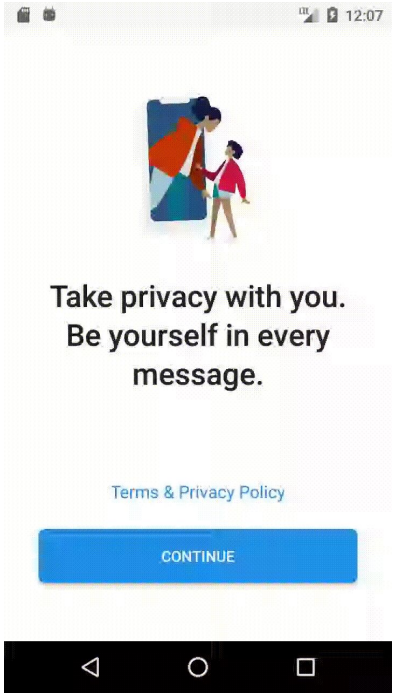


UNIVERSITY OF MARYLAND

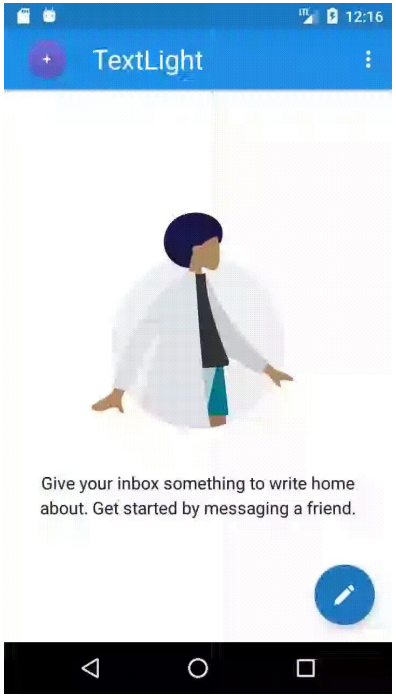
Based on your understanding of end-to-end encryption, please indicate whether you agree or disagree that **hackers who have compromised the TextLight servers** have the following abilities, regardless of their motivation to do so.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Can see that you have sent a message on TextLight, regardless of knowing the content of the message.	○	○	●	○	○
Can see what is in the	○	○	●	○	○

Pre survey | Remote install | Daily chat | Exit survey | Optional interview



Pre survey | Remote install | Daily chat | Exit survey | Optional interview



Pre survey | Remote install | Daily chat | Exit survey | Optional interview

Short texting sessions for 20 days.

Pre survey | Remote install | Daily chat | Exit survey | Optional interview

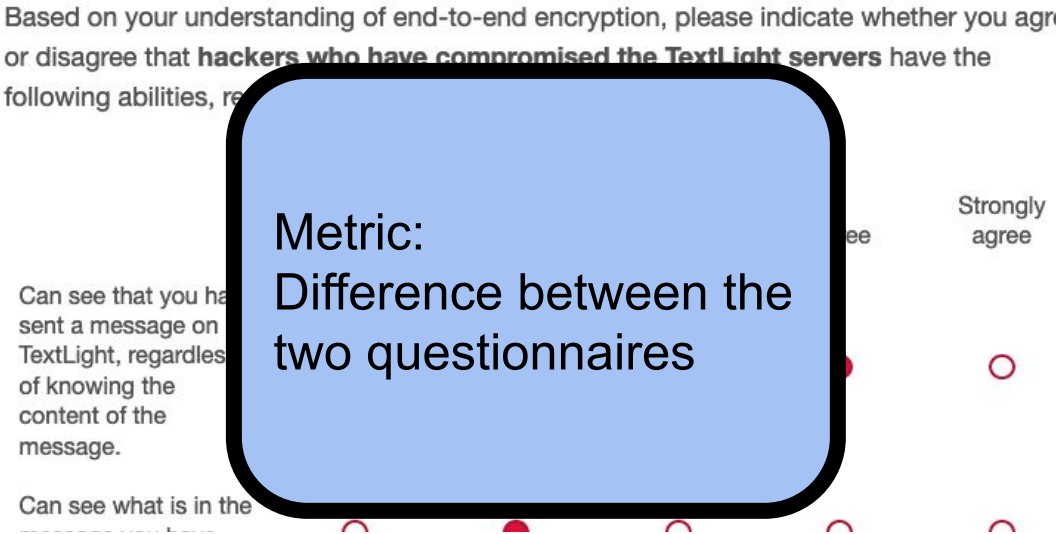
UNIVERSITY OF MARYLAND

Based on your understanding of end-to-end encryption, please indicate whether you agree or disagree that **hackers who have compromised the TextLight servers** have the following abilities, regardless of their motivation to do so.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Can see that you have sent a message on TextLight, regardless of knowing the content of the message.	○	○	○	●	○
Can see what is in the	○	●	○	○	○

Pre survey | Remote install | Daily chat | Exit survey | Optional interview

UNIVERSITY OF MARYLAND



Metric:
Difference between the two questionnaires

Pre survey | Remote install | Daily chat | Exit survey | Optional interview

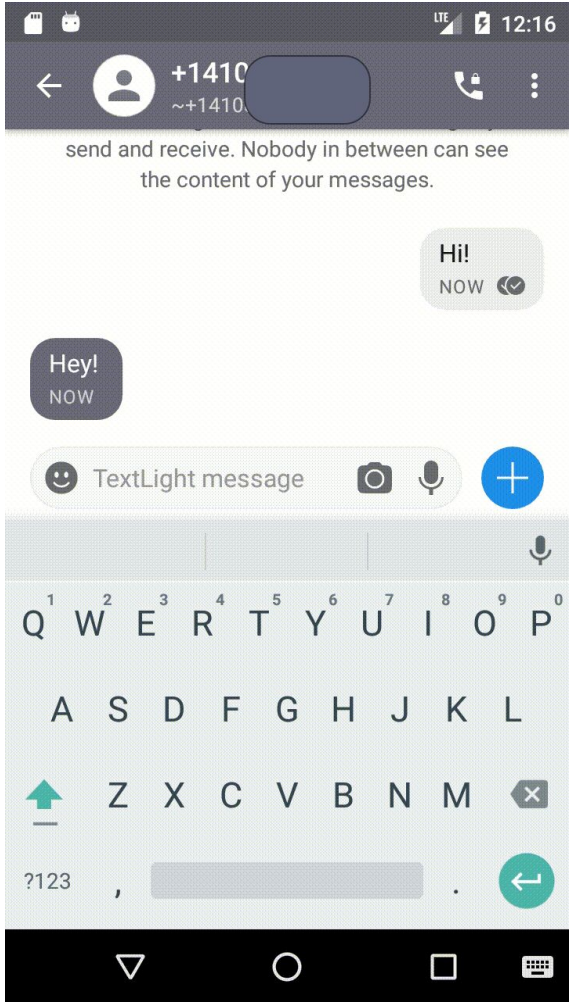
Pre survey | Remote install | Daily chat | Exit survey | Optional interview

Icon made by Becris from flaticon.com

Study 3: Results

- Statistically, there is no improvement between experimental and control groups
  - People knew more than expected
  - In one question, we oversold E2EE (like in the survey study)

- Interviews tell us more



Interviews:

- 10/19 participants were able to generalize the concept
  - " [it protects from] *Probably anyone who would interrupt or interfere in between the messaging, in between where you sent it and someone else received it.*"
- 14/19 knew the unlocked phone adversary was powerful
- 9/19 got at least something wrong
  - " [it protects from] *people … hacking into your phone … from either reading the messages or altering the contents of the message.*"
- 9/19 said they didn't read messages or weren't interested in them.
  - "*I obviously didn't pay a lot of attention to it.*"

Summary

- Mental models of secure communication: not **functional** enough
- Can **small nudges** and user-centered design improve things?
  - Initial qualitative study to identify topics, messages
  - Online study to examine specific messages
  - Longitudinal study to measure real-world effectiveness

- They work well when we **control external factors**.
- **Integration** to applications might need to be more obvious.
  - Perhaps by sacrificing usability a little bit.

Summary

Questions? → akgul@umd.edu

- Mental models of secure communication: not **functional** enough
- Can **small nudges** and user-centered design improve things?
  - Initial qualitative study to identify topics, messages
  - Online study to examine specific messages
  - Longitudinal study to measure real-world effectiveness

- They work well when we **control external factors**.
- **Integration** to applications might need to be more obvious.
  - Perhaps by sacrificing usability a little bit.

References

1. W. Bai, D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek. An inconvenient trust: User attitudes toward security and usability tradeoffs for key-directory encryption systems. In Twelfth Symposium on Usable Privacy and Security (SOUPS 2016), pages 113–130, Denver, CO, June 2016. USENIX Association.
2. W. Bai, D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek. Balancing security and usability in encrypted email. IEEE Internet Computing, 21(3):30–38, May 2017.
3. Wei Bai, Michael Pearson, Patrick Gage Kelley, and Michelle L. Mazurek. Improving Non-Experts' Understanding of End-to-End Encryption: An Exploratory Study. In IEEE 5th European Workshop on Usable Security (EuroUSEC), 2020.
4. W. Diffie and M. E. Hellman. New directions in cryptography. Information Theory, IEEE Transactions on, 22(6):644–654, Nov 1976.
5. S. Fahl, M. Harbach, T. Muders, and M. Smith. Confidentiality as a Service – usable security for the cloud. In Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on, pages 153–162, June 2012.
6. S. Fahl, M. Harbach, T. Muders, M. Smith, and U. Sander. Helping johnny 2.0 to encrypt his facebook conversations. In Proceedings of the Eighth Symposium on Usable Privacy and Security, SOUPS '12, pages 11:1–11:17. ACM, 2012.
7. S. L. Garfinkel and R. C. Miller. Johnny 2: a user test of key continuity management with S/MIME and Outlook Express. In Proceedings of the 2005 Symposium on Usable Privacy and Security, SOUPS '05, pages 13–24. ACM, 2005.
8. S. Gaw, E. W. Felten, and P. Fernandez-Kelly. Secrecy, flagging, and paranoia: Adoption criteria in encrypted email. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06, pages 591–600, New York, NY, USA, 2006. ACM.

References

9. S. E. McGregor, P. Charters, T. Holliday, and F. Roesner. Investigating the Computer Security Practices and Needs of Journalists. In 24th USENIX Security Symposium (USENIX Security 15), pages 399–414. USENIX Association, 2015.
10. M. S. Melara, A. Blankstein, J. Bonneau, E. W. Felten, and M. J. Freedman. CONIKS: Bringing key transparency to end users. In 24th USENIX Security Symposium (USENIX Security 15), pages 383–398. USENIX Association, Aug. 2015.
11. S. Ruoti, J. Anderson, S. Heidbrink, M. O'Neill, E. Vaziripour, J. Wu, D. Zappala, and K. Seamons. "We're on the same page": A usability study of secure email using pairs of novice users. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '16, pages 4298–4308, New York, NY, USA, 2016. ACM.
12. S. Ruoti, N. Kim, B. Ben, T. van der Horst, and K. Seamons. Confused Johnny: when automatic encryption leads to confusion and mistakes. In Proceedings of the Ninth Symposium on Usable Privacy and Securit, SOUPS '13, pages 5:1–5:12. ACM, July 2013.
13. M. D. Ryan. Enhanced certificate transparency and end-to-end encrypted mail. In 21st Annual Network and Distributed System Security Symposium, NDSS'14, 2014.
14. S. Sheng, L. Broderick, C. A. Koranda, and J. J. Hyland. Why Johnny still can't encrypt: evaluating the usability of email encryption software. In Proceedings of the Second Symposium on Usable Privacy and Security, SOUPS '06, 2006.
15. D. J. Solove. 'I've got nothing to hide' and other misunderstandings of privacy. San Diego Law Review, 44:745, 2007.

References

16. W. Tong, S. Gold, S. Gichohi, M. Roman, and J. Frankle. Why King George III can encrypt. http://randomwalker.info/teaching/spring-2014-privacy-technologies/king-george-iiiencrypt.pdf, 2014.
17. A. Whitten and J. D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In Proceedings of the 8th Conference on USENIX Security Symposium - Volume 8, SSYM'99, pages 14–14, 1999.
18. R. Abu-Salma, M. A. Sasse, J. Bonneau, A. Danilova, A. Naiakshina, and M. Smith. Obstacles to the adoption of secure communication tools. In 2017 IEEE Symposium on Security and Privacy (SP), pages 137–153, San Jose, CA, May 2017. IEEE Computer Society.
19. F. Asgharpour, D. Liu, and L. J. Camp. Mental models of security risks. In Proceedings of the 11th International Conference on Financial Cryptography and 1st International Conference on Usable Security, FC'07/USEC'07, pages 367–377, Berlin, Heidelberg, 2007. Springer-Verlag.
20. S. Dechand, A. Naiakshina, A. Danilova, and M. Smith. In encryption we don't trust: The effect of end-to-end encryption to the masses on user perception. In 2019 IEEE European Symposium on Security and Privacy (EuroS&P), pages 401–415, Stockholm, Sweden, June 2019. IEEE Computer Society.
21. A. Demjaha, J. Spring, I. Becker, S. Parkin, and A. Sasse. Metaphors considered harmful? an exploratory study of the effectiveness of functional metaphors for end-to-end encryption. In Workshop on Usable Security. Internet Society, 2018.
22. Electronic Frontier Foundation. Secure Messaging Scorecard, 2016. https://www.eff.org/node/82654.
23. N. Gerber, V. Zimmermann, B. Henhapl, S. Emeröz, and M. Volkamer. Finally johnny can encrypt: But does this make him feel more secure? In Proceedings of the 13th International Conference on Availability, Reliability and Security, ARES 2018, pages 11:1–11:10, New York, NY, USA, 2018. ACM.

References

24. J. R. C. Nurse, S. Creese, M. Goldsmith, and K. Lamberts. Trustworthy and effective communication of cybersecurity risks: A review. In 2011 1st Workshop on Socio-Technical Aspects in Security and Trust (STAST), pages 60–68, Sep. 2011.
25. S. Schröder, M. Huber, D. Wind, and C. Rottermanner. When Signal Hits the Fan: On the Usability and Security of State-of-the-Art Secure Mobile Messaging. In European Workshop on Usable Security (EuroUSEC), Darmstadt, Germany, 2016. Internet Society.
26. J. Tan, L. Bauer, J. Bonneau, L. F. Cranor, J. Thomas, and B. Ur. Can unicorns help users compare crypto key fingerprints? In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17, pages 3787–3798, New York, NY, USA, 2017. ACM.
27. E. Vaziripour, J. Wu, M. O'Neill, D. Metro, J. Cockrell, T. Moffett, J. Whitehead, N. Bonner, K. Seamons, and D. Zappala. Action needed! helping users find and complete the authentication ceremony in signal. In Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018), pages 47–62, Baltimore, MD, August 2018. USENIX Association.
28. E. Vaziripour, J. Wu, M. O'Neill, J. Whitehead, S. Heidbrink, K. Seamons, and D. Zappala. 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), pages 29–47, Santa Clara, CA, July 2017. USENIX Association.
29. J. Warshaw, N. Taft, and A. Woodruff. Intuitions, analytics, and killing ants: Inference literacy of high school-educated adults in the US. In Twelfth Symposium on Usable Privacy and Security (SOUPS 2016), pages 271–285, Denver, CO, June 2016. USENIX Association.

References

30. J. Wu, C. Gattrell, D. Howard, J. Tyler, E. Vaziripour, D. Zappala, and K. Seamons. "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), pages 137–154, Santa Clara, CA, August 2019. USENIX Association.
31. J. Wu and D. Zappala. When is a tree really a truck? Exploring mental models of encryption. In Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018), pages 395–409, Baltimore, MD, August 2018. USENIX Association.
32. K. Krombholz, K. Busse, K. Pfeffer, M. Smith, and E. Zezschwitz, "'If HTTPS were secure, I wouldnt need 2FA': End user and administrator mental models of HTTPS," in IEEE Symposium on Security and Privacy, May 2019.