

Social Cybersecurity: Reshaping Security Through An Empirical Understanding of Human Social Behavior

CMU-HCII-17-100

May 2017

Sauvik Das

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Thesis Committee

Jason I. Hong (chair) [HCII, CMU]
Laura A. Dabbish (co-chair) [HCII, CMU]
Jeffrey P. Bigham [HCII, CMU]
J.D. Tygar [University of California, Berkeley]

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

Copyright © 2017 Sauvik Das, All Rights Reserved.

This research was supported in part by the National Defense Science and Engineering Graduate Fellowship, the Qualcomm Innovation Fellowship, Carnegie Mellon University, as well as by the National Science Foundation under award number 1347186. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or any other sponsors.

Keywords: Cybersecurity, Usable Privacy and Security, Social Psychology, HCI, Data Science, Ubiquitous Computing, Social Computing, CSCW

Abstract

Despite substantial effort made by the usable security community at facilitating the use of recommended security systems and behaviors, much security advice is ignored and many security systems are underutilized. I argue that this disconnect can partially be explained by the fact that security behaviors have myriad unaccounted for social consequences. For example, by using two-factor authentication, one might be perceived as “paranoid”. By encrypting an e-mail correspondence, one might be perceived as having something to hide. Yet, to date, little theoretical work in usable security has applied theory from social psychology to understand how these social consequences affect people’s security behaviors. Likewise, little systems work in usable security has taken social factors into consideration.

To bridge these gaps in literature and practice, I begin to build a theory of social cybersecurity and apply those theoretical insights to create systems that encourage better cybersecurity behaviors. First, through a series of interviews, surveys and a large-scale analysis of how security tools diffuse through the social networks of 1.5 million Facebook users, I empirically model how social influences affect the adoption of security behaviors and systems. In so doing, I provide some of the first direct evidence that security behaviors are strongly driven by social influence, and that the design of a security system strongly influences its potential for social spread. Specifically, security systems that are more *observable*, *inclusive*, and *stewarded* are positively affected by social influence, while those that are not are negatively affected by social influence.

Based on these empirical results, I put forth two prescriptions: (i) creating socially grounded interface “nudges” that encourage better cybersecurity behaviors, and (ii) designing new, more socially intelligent end-user facing security systems. As an example of a social “nudge”, I designed a notification that informs Facebook users that their friends use optional security systems to protect their own accounts. In an experimental evaluation with 50,000 Facebook users, I found that this social notification was significantly more effective than a non-social control notification at attracting clicks to improve account security and in motivating the adoption of promoted, optional security tools. As an example of a socially intelligent cybersecurity system, I designed Thumprint: an inclusive authentication system that authenticates and identifies individual group members of a small, local group through a single, shared secret knock. Through my evaluations, I found that Thumprint is resilient to casual but motivated adversaries and that it can reliably differentiate multiple group members who share the same secret knock. Taken together, these systems point towards a future of socially intelligent cybersecurity that encourages better security behaviors. I conclude with a set of descriptive and prescriptive takeaways, as well as a set of open problems for future work.

Concretely, this thesis provides the following contributions: (i) an initial theory of social cybersecurity, developed from both observational and experimental work, that explains how social influences affect security behaviors; (ii) a set of design recommendations for creating socially intelligent security systems that encourage better cybersecurity behaviors; (iii) the design, implementation and comprehensive evaluation of two such systems that leverage these design recommendations; and (iv) a reflection on how the insights uncovered in this work can be utilized alongside broader design considerations in HCI, security and design to create an infrastructure of useful, usable and socially intelligent cybersecurity systems.

Acknowledgements

If you would have asked me, when I began my Ph.D., “What does it take to thrive in academia?”, I would have guessed intelligence and hard work. It didn’t take long for me to realize that intelligence and hard work, while important, are essentially constant across everyone who chooses this path. Implicit in my understanding of academia was a (faulty) assumption that success was a matter of individual effort. In reality, I’ve learned that at least my personal success in academia has been a function of my support network of friends, family, colleagues and mentors.

The work I’ve written in this document is the culmination of six years of careful learning from others. Learning, for example, how to pick research questions, how to design experiments, how to build systems, how to evaluate those systems, how to make grounded inferences from a pool of independent results, how to be skeptical of my own work, and, importantly, how to be okay with failure and rejection. Without my support network, I could not have learned these essential skills.

I owe a great deal to my advisers, Jason Hong and Laura Dabbish. Jason took me on as a student just as soon as I entered the HCII and has been invaluable to my growth and development ever since. I am fortunate to have found an adviser, in Jason, who perfectly matched my own personality: he offered me the freedom and leave to make my own mistakes while providing me with the guidance and mentorship to ensure that I learned from those mistakes and developed into a good researcher. Laura, who mentored me first as a project co-adviser and later as a full co-adviser brought with her a level of enthusiasm and insight of human behavior that helped me push through all doubts. I was convinced that I would succeed in my dissertation work owing in no small part to Laura’s guidance.

I’ve also had many mentors outside of my advisers. To my committee member, Jeffrey Bigham (CMU), I offer my most sincere “Yo.” Your approach to academia and your job is one that I’ve always admired and one that I sincerely hope to emulate as I take my own next step. To my external committee member, Doug Tygar (Berkeley), I am grateful that you took the time to help shape and strengthen my work. I am fortunate to have had your guidance and consider it a great honor that you took the time. I hope we can have many more interesting conversations in the future.

I am also thankful for the many mentors with whom I have personally worked and who helped shape my approach to research. These include: Steven Dow (UCSD), Adrian Perrig (ETH), Mark Riedl (Georgia Tech), Tom Zimmermann (Microsoft Research), Chris Harrison (CMU), Adam Kramer (Facebook), Stuart Schechter (Samsung), Sebastian Schnorf (Google), Alexander de Luca (Google) and Koji Yatani (University of Tokyo).

I’ve also been fortunate to have had many helpful conversations about research and my career with others, including Lorrie Cranor (CMU), Amy Bruckman (Georgia Tech), Justin Cranshaw (Microsoft Research), Andy Ko (University of Washington), Moira Burke (Facebook), Johan Ugander (Stanford), Michael Bernstein (Stanford) and Eric Gilbert (University of Michigan). Thank you all!

I have worked with many amazing colleagues in pursuit of the work in this document. I say with little hesitation that this work would not have been possible without the helpful contributions of: Tiffany Hyun-Jin Kim, Melissa Luu-Van, Joanne Lo, David Lu, Taehoon Lee and Alex Sciuto.

In addition to all of these mentors and colleagues, I am indebted to the many friends I have made along the way. Life at CMU would be so much less fulfilling without the friendship of my old friend Mike, my wonderful cohort (Erik, Kelly, Nesra, Robert, Ryan, Samantha, Sarah), the CHIMPS Lab (Jason Wiese, Eiji, Shah, Annabel), the clique (Dan, Tati, Jeff Rz., Beka) and the many, many other friends I’ve made along the way (including Ian, Min, Stephen, Walter, Chloe, Eliane, Ruogu, Kerry, Yanjin, Chris M., Caitlin, Dave, Nikola, Brandon, Jenny, Anthony, Gierad, Aurora).

To Anna: thank you for supporting me through the home stretch. You’ve made my time in Pittsburgh and the HCII so much more than just grad school. You’ve made it special.

Finally, thank you to my family: Samir, Sangeeta, Swagata, Agassi, Nairi, Uma, Saurav, Caroline, Amira, Rayan, Sandeep, Amisha and Ariana. You’ve all endured many years of my being selfish as I pursued

this path. I truly could not have made it this far without you. Words cannot express how much you all mean to me.

Throughout my Ph.D., I have been very fortunate to work with some of the brightest minds in Usable Privacy and Security, HCI, Design, Data Science and Social Psychology across a multitude of organizations and institutions. It would be beyond my capacity to list all of the amazing people with whom I have worked and who have touched my life and career. With the guidance and support of these friends and mentors, I can finally say that I have made some headway. The six years leading to this point have been difficult, but also intensely fun and enlightening.

I am eternally grateful to everyone who has fueled and supported my growth in this journey.

TABLE OF CONTENTS

ABSTRACT	3
CHAPTER 1: INTRODUCTION AND MOTIVATION	10
THESIS STATEMENT	10
MOTIVATION	10
CHAPTER 2: BACKGROUND AND RELATED WORK	14
WHAT INHIBITS GOOD CYBERSECURITY BEHAVIORS?	14
HOW CAN SOCIAL INFLUENCE ENCOURAGE BETTER CYBERSECURITY BEHAVIORS?	15
WHAT ARE THE GAPS IN OUR UNDERSTANDING? WHAT ARE THE OPPORTUNITIES?	17
CHAPTER 3: A TYPOLOGY OF HOW SOCIAL INFLUENCE AFFECTS SECURITY BEHAVIORS	18
SUMMARY	18
MOTIVATION	18
METHODOLOGY	19
RESULTS	21
DISCUSSION	27
CHAPTER 4: UNDERSTANDING HOW SECURITY INFORMATION IS COMMUNICATED	29
SUMMARY	29
MOTIVATION	29
METHODOLOGY	29
RESULTS	30
DISCUSSION	38
CHAPTER 5: HOW SOCIAL INFLUENCES AFFECT SECURITY TOOL DIFFUSION	39
SUMMARY	39
MOTIVATION	39
HYPOTHESES	40
METHODOLOGY	41
RESULTS	43
DISCUSSION	49
LEVERAGING SOCIAL INFLUENCE TO IMPROVE END-USER SECURITY	50
CHAPTER 6: INCREASING SECURITY SENSITIVITY WITH SOCIAL PROOF	53
SUMMARY	53
MOTIVATION	53
EXPERIMENT	54
HYPOTHESES	59
RESULTS	60
FOLLOW-UP STUDY WITH SURVEY TO GAUGE SENTIMENT AND AWARENESS	63
DISCUSSION	65
CHAPTER 7: THUMPRINT	67
SUMMARY	67
MOTIVATION	68
SYSTEM DESIGN	69
FEASIBILITY EVALUATION	73
PROCEDURE	73
RESULTS	74
CONSISTENCY AND SECURITY EVALUATION	75

PROCEDURE	75
RESULTS	76
DISCUSSION	78
CHAPTER 8: DISCUSSION & CONCLUSIONS	80
SUMMARY	80
TAKE-AWAYS	81
OPEN PROBLEMS AND THE FUTURE OF SOCIAL CYBERSECURITY	83
CONCLUSION	85
REFERENCES	87

List of Tables

Table 1. Interview participant demographics, occupations, use of authentication on their mobile phones, as well as social media usage. _____	20
Table 2. Inter-coder rating agreement on 20% of the behavior change excerpts. _____	21
Table 3. Social triggers for behavior change derived from our iterative open coding process. _____	22
Table 4. Inter-coder agreement of codes on a 20% random sample of communication excerpts. _____	30
Table 5. Conversation catalysts derived from our iterative open coding process. _____	31
Table 6. Conversation goals derived from our iterative open coding process. _____	33
Table 7. Co-frequency of catalysts for conversations about security and privacy (rows) and reasons for starting the conversation (columns). _____	34
Table 8. The most frequent conversations about security and privacy, based on the catalyst and content. _____	34
Table 9. Collected tool descriptions. These variables were all collected per individual. _____	42
Table 10. Coefficients for the three logistic regressions relating social proof variables (bolded, at the bottom), to use of login approvals (left), login notifications (middle) and trusted contacts (right). All coefficients are normalized. _____	44
Table 11. Exposed condition prerequisites for each security tool. For example, if a user is “exposed” at E3 for login approvals, at least 1.3% of her friends must have adopted login approvals at the time of data collection. _____	46
Table 12. Chi square significance tests for the difference in adoption rate between exposed and unexposed individuals across all exposure conditions and all security features. All differences significant, $p < 2e-16$. _____	47
Table 13. Prompt text in announcement across all 8 experimental groups. Some social groups have templates that are filled in with either the number or percentage of a user’s security feature-using friends. _____	56
Table 14. Collected feature descriptions and distributions for the $n=50,000$ people in our sample. † Approximate values. _____	58
Table 15. Clicks and adoptions by experimental conditions. “N” represents the number of users who viewed the announcement. “ST” stands for short term, and “LT” stands for long term. These values are strictly descriptive. Statistical tests used and significance is mentioned where relevant in the text. _____	60
Table 16. Coefficients for the three regressions predicting clicks, feature adoptions up to a week after the experiment, and feature adoption up to 5 months after the experiment. Bolded coefficients are of interest. _____	61
Table 17. Number of survey responses per solicitation method (rows) and experimental group (columns). _____	64
Table 18. Coefficients for the two proportional-odds logistic regressions predicting agreement with the trustworthy and protection statements. _____	64
Table 19. Features extracted for every thumbprint, drawn from recommendations in prior work in sensing techniques. In total, 1020 features are extracted, though the feature space is dramatically reduced in later steps to avoid overfitting. _____	71
Table 20. Study 2 flow for each participant. The columns represent the six thumbprints selected from Study 1. Cell values with “main” refer to thumbprints participants learned in session 1 and replicated in session 2. Other cell values refer to thumbprints replicated as adversaries. _____	75

List of Figures

- Figure 1.** An outline of my thesis. I start by compiling prior work in social psychology and usable security, outlining pertinent interconnections (Chapter 2). I next present a series of exploratory analyses to model how social psychological processes affect security behaviors, synthesizing two prescriptions (Chapters 3-5). I then create and evaluate illustrative examples of these prescriptions (Chapters 6-7). Finally, I reflect on a set of key takeaways and implications from all of the work (Chapter 8). _____ 13
- Figure 2.** The security sensitivity stack is a theoretical framework to help understand why people ignore security advice or avoid using security tools and best practices. The basic argument is that people either are unaware of threats and tools, not motivated to act on security concerns, or do not know how to use security tools and best practices. _____ 15
- Figure 3.** The number of times each social trigger for behavior change reported by our sample raised any of the three parts of the security sensitivity stack: awareness, motivation, or knowledge. _____ 23
- Figure 4.** Coefficients for the three logistic regressions relating the number of diverse social contexts variable to use of each security tool, with 95% confidence intervals. All coefficients significant, $p < 2e-16$. _____ 44
- Figure 5.** Histogram of percent of friends who use login approvals (left), login notifications (middle) and trusted contacts (right). Colors represent up to what exposed conditions users with x% of tool-adopting friends would be considered “exposed” in the analysis _____ 46
- Figure 6.** Feature adoption rates, plotted for each security feature for each exposure condition, for both exposed and unexposed individuals. Exposed feature adoption rates are plotted as red circles, and unexposed feature adoption rates are plotted as blue triangles. _____ 47
- Figure 7.** Differences in adoption rate between the exposed and unexposed for all three features across all exposure conditions. Values above the dashed horizontal line signify that those who were exposed had a higher adoption rate than the unexposed. All differences were significant at $p < 2e-16$ _____ 48
- Figure 8.** Users saw our announcements on top of their newsfeeds, as shown above, up to 3 times. ____ 55
- Figure 9.** Screenshots of each of the social framings and the control announcements in our experiment. In total we had 8 announcements. Not pictured is the Raw %, Over # and Only % announcements, but they look similar to their counterparts pictured above. _____ 57
- Figure 10.** With Thumprint, groups of users learn a single, shared secret knock that they enter on a surface instrumented with (or containing) an accelerometer and microphone (here, a smartphone) in order to authenticate. _____ 67
- Figure 11.** With Thumprint, users enter secret knocks on an instrumented sensor surface (A) from which a variety of time and frequency domain features are extracted (B). These readings are projected onto a reduced feature space, where each authentication attempt is compared against previously learned thumprint expressions from group members (C). If a match, Thumprint will provide access by regulating an end-point such as an electronic lock (D). _____ 70
- Figure 12.** To authenticate, an unlabeled feature vector is transformed into the reduced feature space and then its distance to nearby training clusters is calculated. In this case, the unlabeled attempt would not be authenticated because it is too far from candidate clusters. _____ 72
- Figure 13.** Screenshots of the app in which participants entered preset (left) and custom (middle) thumprints. The right most figure shows how participants actually used the application interface. ____ 73
- Figure 14.** Mean feature vector difference (along with 95% confidence intervals) for user testing attempts (relative to their own training data and other group member training data), as well as outsider attempts. _____ 74
- Figure 15.** Acceptance rate as a function of feature vector difference. The black vertical line is where 100% of user attempts are accepted, and the blue dashed line is where >0% of outsider attempts are first accepted. _____ 74
- Figure 16.** Mean feature vector difference (along with 95% confidence intervals) for T1-T3 across authentic and adversarial attempts. User testing data was collected one day after the training data. _ 76
- Figure 17.** Acceptance rate as a function of minimum acceptable threshold across all thumprints. There is no threshold value to perfectly distinguish authentic attempts from adversarial attempts, but threshold values between 0.4 and 0.5 yield high true positives and low false positives. _____ 77

Chapter 1: Introduction And Motivation

Thesis Statement

Cybersecurity helps realize the full potential of computing. Without authentication and encryption, for example, few would use digital wallets, social media or even e-mail. The struggle of security is to realize this potential without imposing too steep a cost. Yet, for the average non-expert, security is just that: too costly, in terms of at least time and effort. However, despite substantial improvements to usability of security systems and behaviors, much security advice goes ignored and many useful security systems remain underutilized. I argue that this disconnect can partially be explained by the fact that security behaviors have myriad unaccounted for social consequences. In other words, there's an additional largely unconsidered cost to good security behaviors: social capital. For example, by using two-factor authentication, one might be perceived as "paranoid". By encrypting an e-mail correspondence, one might be perceived as having something to hide. It is unsurprising, therefore, that for many laypeople, even usable security tools are begrudgingly tolerated if not altogether subverted. Accordingly, in this thesis, I present evidence in support of the following claim: **Social influences strongly affect cybersecurity behaviors, and it is possible to encourage better cybersecurity behaviors by designing security systems that are more social.** In support of this statement, I build an initial theory of how social influences affect cybersecurity behaviors, distill these theoretical insights into a set of broad design recommendations, and then implement and evaluate two systems that point to a future of social intelligent cybersecurity.

Motivation

In early 2013, the Associated Press's Twitter account was compromised through a password phishing scheme, and erroneously tweeted that President Obama was injured in a bombing [113]. In response, stock prices plummeted [78], adversely affecting thousands. Moreover, this incident could have been easily prevented with the use of two-factor authentication: A security tool, available at that time, that requires entry of a random code generated on one's phone in addition to a password when authenticating [63]. This incident is just one example of how the underutilization of available security tools and recommended security behaviors remains a large, outstanding problem that begs the question: *How can we design systems that encourage better cybersecurity behaviors?*

Engaging with this question is *important* because cybersecurity unlocks the full potential of computing in society. Without authentication and encryption, for example, few would use digital wallets, social media or even e-mail. The struggle of security and privacy is to realize this potential without imposing too steep a cost. Yet, for the average person, today, security is just that: costly, in terms of time and effort. It is unsurprising, therefore, that for many people, security is begrudgingly tolerated if not altogether subverted, as illustrated by the previous example.

Engaging with this question is *urgent* because as computing encompasses more of our lives, we are tasked with making increasingly more and increasingly important security decisions. Today, a security breach might compromise sensitive data about our finances and schedules as well as deeply personal data about our health, communications, and interests. Tomorrow, as we enter an era of pervasive smart things, that breach might compromise access to our homes (through IoT devices like smart locks), vehicles (through smart cars) and bodies (through wearables and networked medical devices). In other words, we are rapidly immersing ourselves in an environment where cybersecurity breaches will start having physical safety consequences.

Accordingly, there have been many efforts in both industry and academia seeking to address the underutilization of security systems and behaviors. Among academics, there has been a deep focus on improving the usability of existing security systems (e.g., [62,64,93,111]), inventing new security systems that are made to be usable (e.g., [27,51,53,61,96]), as well as better communicating information about security risks and options to counter-act those risks through warnings and notifications (e.g., [14,38,39,42,43]). Likewise, end-user facing online companies like Facebook and Google are investing a lot in the construction of security and privacy tools to provide people with a multitude of options to secure their accounts and protect their data. For example, Facebook offers a suite of optional, end-user facing security tools such as Login Notifications (e-mail/SMS notifications of login attempts), Login Approvals (two-factor authentication) and Trusted Contacts (enlisting

friends to help with account recovery), while Google has invested considerable effort in improving reactive and preventative warnings to keep consumers safe from privacy and security threats.

Nevertheless, despite these substantial and important usability improvements, the awareness and adoption of these tools remains relatively low [28,63]. For example, at least in 2013, the adoption rate, among Facebook users, was less than 10% for Login Notifications, less than 2% for Login Approvals and Trusted Contacts [30]. Similarly, a field study on the effectiveness of browser warnings found that as many as 70% of users bypass certain security and privacy warnings [5].

Why does the adoption of recommended security behaviors and systems remain low? Prior work offers a number of reasons. Some prior work suggests that many believe they are in no danger of experiencing a security breach [3] and are *unaware* of existing security threats and the tools available to protect themselves against those threats. Other work suggests that many choose not to use security tools and follow security advice because doing so is often antagonistic towards the immediate goal of end users—a complex password that often requires three attempts to get right *prevents* a user from doing what she actually wants to do: e.g., authenticating into social media. Herley further argues it may even be economically *rational* for users to ignore security advice, as the expected cost over a lifetime of following security advice might actually be higher than the expected loss one would suffer if his account actually was compromised [55]. Thus, many people are *unmotivated* to behave securely. Still others suggest that security tools and behaviors are too difficult to use [92,111], so many people do not have the *knowledge* required to operate them. Taken together, the lack of what I call *security sensitivity*—the awareness of, motivation to use, and knowledge of how to use security tools—is a barrier to increasing the uptake of security tools and the following of security advice.

I argue that one reason security sensitivity remains low among the general populace is that we do not yet understand the *social* processes underlying people’s decisions to communicate about security and adopt security tools. In other words, security behaviors—as any human behavior—should be viewed within the context of a broader sociotechnical system. Indeed, the social psychology literature illustrates that social influence, or our ability to affect other people’s perceptions and behaviors with our words and actions [21], plays a central role in how people behave. This effect applies even in adopting a new technology or idea [21,88]. Rogers’ highly influential *diffusion of innovations* work, for example, has shown that social influence drives technology adoption [88]. Likewise, another popular model for explaining how technology gets adopted and sees widespread use, the Technology Acceptance Model [32], identifies social influence as a key factor in driving new technology adoption.

It stands to reason, therefore, that social influences should affect one’s decision to follow security advice or use a security tool (e.g., like two-factor authentication). Importantly, however, whether this effect is positive or negative is yet to be seen, and a case can be made for either. Indeed, prior work has shown that people who encrypt email can be perceived as “paranoid” by others [45], which, in turn, suggests that the early adopters of encryption may cause others to “disaffiliate” themselves from encryption because they, themselves, don’t want to be seen as “paranoid”. Conversely, the principle of *social proof*—that we look to others for cues on how to behave, especially when we are uncertain [21]—suggests that if people see many examples of others using a security tool, they should be *more inclined* to use the tool themselves. Understanding this tension between disaffiliation and social proof, and especially how it interacts with the particular design of different security tools, might be pivotal in determining people’s *motivation* to use a security tool. Similarly, how people *communicate with each other* about security tools and behaviors might be pivotal in determining people’s *awareness* and *knowledge* of cybersecurity. Accordingly, social processes might significantly impact all layers of the security sensitivity stack—*awareness, motivation, and knowledge*.

Yet, to date, little theoretical work in usable privacy and security has applied social science theory to understand how social processes affect security sensitivity and decision making. In turn, this lack of theoretical insight has precluded systems work that accounts for the social consequences of security system design. Thus, there remains a great but largely untapped opportunity in modeling human

social behaviors within the context of cybersecurity and in creating socially intelligent security systems that have a better understanding of these human social behaviors.

To bridge these gaps in theory and practice, in this thesis, I build an initial theory of how social influences affect cybersecurity behaviors, distill these theoretical insights into a set of broad design recommendations, and then implement and evaluate two systems that point to a future of socially intelligent cybersecurity. Concretely, my thesis is organized around the following contributions:

- *A survey of the existing literature connecting social psychology and cybersecurity.* In Chapter 2, I begin with a comprehensive survey of the existing literature that connects insights from social psychology with applications in cybersecurity.
- *An initial descriptive theory of how social processes affect security behaviors.* In Chapters 3-5, I describe a series of observational investigations into how social influences affect security behaviors. Through a series of interviews and surveys, I construct typologies of social triggers for security behavior changes and of how people come to learn about security from their interactions with others. I complement these findings with a large-scale empirical analysis of how security tools diffuse through the social networks of 1.5 million Facebook users. In so doing, I provide some of the first direct empirical evidence that security behaviors are strongly driven by social influences, and that the design of a security system strongly influences its potential for social spread.
- *A set of design recommendations to create socially intelligent cybersecurity systems.* At the end of Chapter 5, I distill the broad descriptive findings from Chapters 3-5 into a set of prescriptive design recommendations. I suggest that security systems that are more *observable*, *inclusive*, and *stewarded* are positively affected by social influence, while those that are not are negatively affected by social influence. Furthermore, I put forth two prescriptions: (i) creating socially grounded interface “nudges” that encourage better cybersecurity behaviors, and (ii) designing new, more socially intelligent end-user facing security systems.
- *The design, implementation and evaluation of two social cybersecurity systems.* In Chapters 6 and 7, I describe my implementation and evaluation of two social cybersecurity systems, drawing from the design recommendations synthesized at the end of Chapter 5. The first is a social interface nudge that informs Facebook users that their friends use optional security systems in an effort to encourage them to do the same. In a 50,000 user experimental evaluation of these nudges against a non-social control, I find that the social nudges are significantly more effective at attracting attention and in motivating security tool adoption. The second is Thumprint: a socially-inclusive authentication system that authenticates and identifies individual group members of a small, local group (e.g., families, small work teams) through a single, shared secret knock. Through a multi-day, in-person lab study, I find that Thumprint is comparably strong, in authentication strength, to other behavioral biometric authenticators (e.g., keystroke dynamics) but has the added benefit of requiring only a single shared secret in order to differentiate between multiple group members, thus mitigating the need for group members to keep authentication secrets from one another.
- *A reflection on how to create an infrastructure of socially intelligent cybersecurity systems.* Finally, in Chapter 8, I conclude with a broader reflection on how the insights uncovered in my thesis work can be utilized alongside broader design considerations in HCI, security and design to create an infrastructure of useful, usable and socially intelligent cybersecurity systems that encourage better cybersecurity behaviors.

Figure 1 offers an overview of my thesis document. Somewhat uniquely, my contributions span theoretical insights about human behavior drawn from empirical work, to functional systems that

draw on those insights in order to improve the state of end-user facing cybersecurity. Accordingly, this thesis is disciplinarily disparate, broadly relevant to the fields of usable security, human-computer interaction, computer science and social psychology.

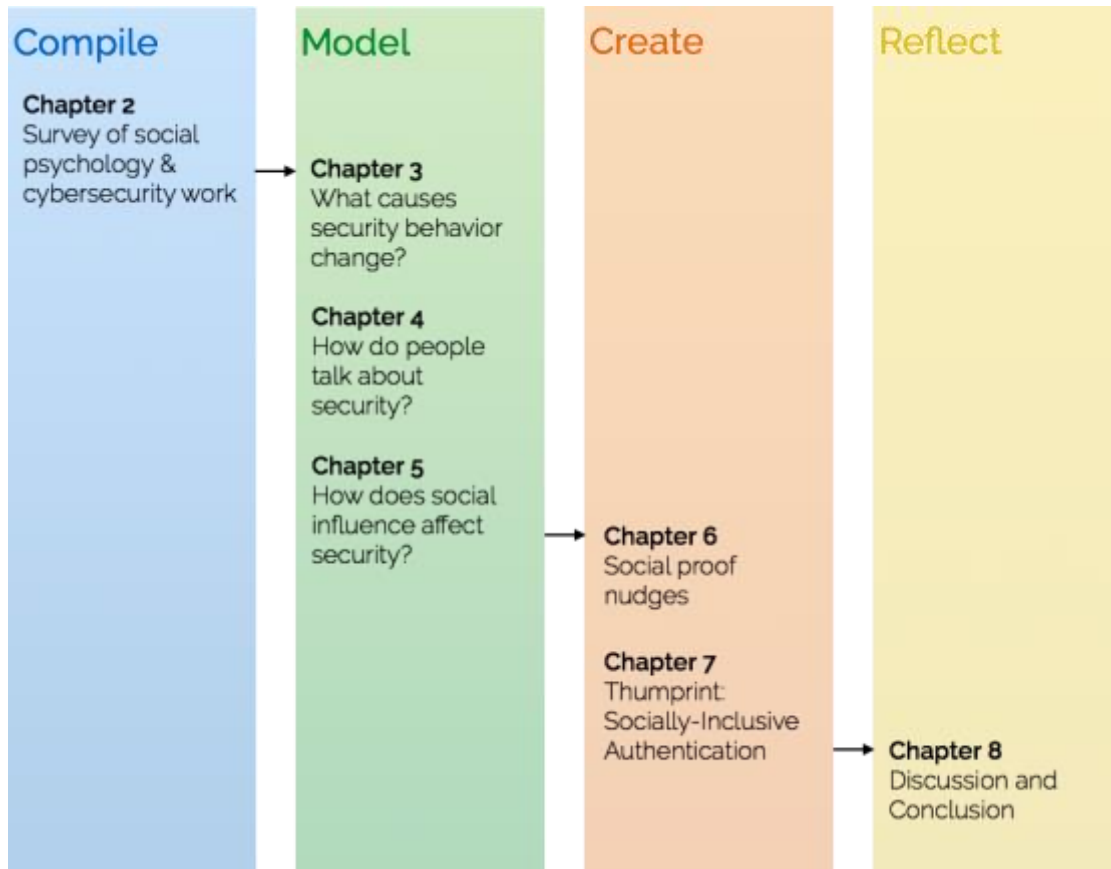


Figure 1. An outline of my thesis. I start by compiling prior work in social psychology and usable security, outlining pertinent interconnections (Chapter 2). I next present a series of exploratory analyses to model how social psychological processes affect security behaviors, synthesizing two prescriptions (Chapters 3-5). I then create and evaluate illustrative examples of these prescriptions (Chapters 6-7). Finally, I reflect on a set of key takeaways and implications from all of the work (Chapter 8).

Chapter 2: Background and Related Work

My thesis draws from and builds upon prior art in social psychology, behavioral economics, ubiquitous computing and cybersecurity. While diverse, the organizing principle behind this disciplinary breadth is its relevance to the principal question of my work: *How can we design systems that encourage better cybersecurity behaviors?*

To that end, I surveyed the background literature across these disciplines based on its pertinence to the answering of more specific, component questions that summarize our current scientific understanding of how to answer that principal question.

What Inhibits Good Cybersecurity Behaviors?

The first step in understanding how to design systems that encourage better cybersecurity behaviors is to understand the barriers to practicing good security behaviors. Prior work in usable privacy and security alludes to at least three reasons underlying why much security advice is ignored and many security tools remain unused: low awareness, motivation, and knowledge.

First, many users have low *awareness* of security threats and the tools available to protect themselves against those threats. For example, a study by Adams and Sasse found that insufficient awareness of security issues caused users to construct their own model of security threats that are often incorrect, resulting in security breaches [4]. Stanton and colleagues found that a lack of awareness of basic security principles even influenced “experts” to make security mistakes, such as using a social security number as a password [100]. Users who are unaware of a threat cannot take measures to avoid the threat, and users who are not cognizant of the tools available to protect themselves from these threats cannot use those tools to actively defend themselves.

Second, users—even those who are aware of security and privacy threats and the preventive tools that combat those threats—often lack the *motivation* to utilize security features to protect themselves [4,38]. This lack of motivation to use security features is not entirely surprising, as stringent security measures are often antagonistic towards the specific goal of the end user at any given moment [36,92]. For example, while a user might want to access her Facebook, a complex password that usually requires three attempts to get right *prevents* her from accessing Facebook for an intolerable amount of time [39].

Negative experiences with or impressions towards security behaviors can also impact motivation. In a survey of over 200 security experts and non-experts, Ion, Reeder and Consolvo found that non-experts tend to practice very different security behaviors than experts believe are important—while experts valued keeping software up-to-date and using password managers, non-experts reported being skeptical of the effectiveness of these behaviors or avoided them because of prior negative experiences with updates [60]. Other work has found that users can have a defeatist attitude towards cybersecurity, believing that if an attacker wanted to access their data they would irrespective of any counter-measures taken [82,83,108].

Low motivation may also be symptomatic of a deeper root cause—that many security threats remain abstract to most individuals [4,57,84]: e.g., Bob may know, conceptually, that there are security risks to using the same simple password across accounts, but does not believe that he is, himself, in danger of experiencing a security breach. Additionally, Herley argues that this perspective may be economically rational, as the expected cost, in monetized time, of a lifetime of following security advice might actually be higher than the expected loss a user would suffer if his account actually was compromised [55]. Finally, the benefits of security features are often invisible, as users are often not cognizant of the *absence* of a breach that otherwise would have occurred without the use of a security or privacy tool. In all, it is unsurprising that many users lack the motivation to explicitly use security tools: to do so would mean to incur a frustrating complication to everyday interactions in order to prevent an unlikely threat with little way to know whether the security tool was actually effective. Beauteamente, Sasse and Bonham frame this broad motivation problem economically as the “compliance budget”—if security costs are too high relative to perceived benefits, “compliance” with security policies is unlikely [10]. More generally, users often reject the use of security and privacy tools when they expect or experience them to be weighty [4,45,59,92].



Figure 2. The security sensitivity stack is a theoretical framework to help understand why people ignore security advice or avoid using security tools and best practices. The basic argument is that people either are unaware of threats and tools, not motivated to act on security concerns, or do not know how to use security tools and best practices.

The third inhibitory barrier to good security behaviors is that users may have low *knowledge* of when, why and how to properly practice good security behaviors. Security tools are often too complex to operate for even aware and motivated end-users, suggesting that users often do not have the specialized *knowledge* to actually utilize security tools [111]. Indeed, there is a wide gulf of execution for most security features for most users. For example, many users cannot distinguish legitimate vs. fraudulent URLs, nor forged vs. legitimate email headers [34]. Another study revealed how security features in Windows XP, Internet Explorer, Outlook Express, and Word applications are difficult for lay users to navigate [44]. Wash found that many people hold “folk” models of computer security that are often misguided, and use these incorrect models to justify ignoring security advice [108]. Knowledge is perhaps the most widely acknowledged and addressed inhibitory barrier to good security behaviors—indeed, an argument can be made that the fundamental goal of usable security, to date, has been to lower the knowledge barrier to practicing good security behaviors.

In sum, prior work in usable security suggests that there are at least three large obstacles inhibiting the widespread use of security tools (see Figure 2): the *awareness* of security threats and tools, the *motivation* to use security tools, and the *knowledge* of how to use security tools. Throughout this document, I refer to this layered stack as *security sensitivity* for ease of discussion, as it encapsulates how likely a user is to seek information about and use security tools. Note, however, that the concept of security sensitivity is not a contribution of this dissertation: prior work has alluded to such a stack in security specifically [38], and in the adoption of technology more generally [33,88].

How Can Social Influence Encourage Better Cybersecurity Behaviors?

Efforts have been made at improving all parts of the security sensitivity stack—for example, through games for security education [98], browser extensions to make people more aware of phish [114], more effective user interfaces for security tools [35] and simpler ways to authenticate [27]. The related work mentioned here is just a sampling of the enormous effort put forth by the usable security community at creating systems, tools and behavioral interventions to make cybersecurity simpler, faster and easier for end-users. Security sensitivity, nevertheless, remains low.

Prior work in cognitive psychology highlights the influential nature of social proof in driving human behavior. In particular, much prior work has demonstrated the potency of the concept of “social proof”—or our tendency to look to others for examples of how to act when uncertain [21,24]. For example, Milgram, Bickman, and Berkowitz [76] demonstrated the social proof principle when they showed that simply getting a small crowd of people—the more, the better—to look up at the sky on a busy sidewalk caused others to do the same. Still other work has shown how social interventions can

be powerfully effective at driving human behavior: for example, at reducing household energy consumption by showing people their neighbors' reduced energy consumption [95], reducing hotel guests' wasteful use of towels by showing them that previous patrons chose to be less wasteful [48], and even in eliminating young children's phobia of dogs by showing them film clips of other children playing with dogs [8].

Other work highlights the significant effect of social processes in the adoption of technology, specifically. For example, in his seminal work on the diffusion of innovations, Rogers claimed that new technology gets widely adopted through a process by which it is communicated through members of a social network [88]. Rogers argued that primarily *subjective perceptions*, not empirical fact, get communicated through social channels, and that these perceptions are key to the success of an innovation spreading. He further outlined that preventative innovations—or innovations, like security and privacy tools, that prevent undesirable outcomes from happening in the future—typically have low adoption rates, probably because of their lack of *observability*, or the invisibility of their use and benefits. More recent studies on online platforms such as Facebook have similarly alluded to the potency of social proof. Kramer [70] showed that users were more likely to share emotional content matching the valence of content shared by friends in the past few days, and Burke and colleagues [17] showed that social learning plays a substantial role in influencing how newcomers to Facebook use the platform. Notably, Bond and colleagues [12] found that simply showing people that their Facebook friends voted was sufficient to increase voter turnout in the 2010 U.S. Congressional elections.

Taken together, the background literature suggests that social influence strongly affects people's behaviors and decisions; likely, also their security-related behaviors and decisions. And, indeed, prior work *has* alluded to the importance of social processes in *raising* security sensitivity. For example, DiGioia and Dourish [35] suggested that “social navigation”—or people's inclination to look for cues on how to act—can be used to raise users' security sensitivity by showing them other users' actions in context. Rader et al.'s study on stories as informal lessons about security suggests that storytelling increases awareness of and motivation to guard against security threats [81]. On the other hand, social processes can also *lower* security sensitivity and/or encourage unsafe practices. For example, Singh et al. outlined the common practice of sharing passwords and PINs [99]. Gaw et al. [45] found that many people believed that use of security tools was an indication of paranoia, unless the user had an obvious reason for doing so. If there is a stigma of paranoia attached to using security features, then it is possible that *social influence can also work against security sensitivity* (e.g., “only paranoid people encrypt their e-mail, and I'm not paranoid”).

In fact, because security tool usage is often invisible, rarely communicated, and generally undesired [15, 24], it may be that social processes, left unchecked, work against security sensitivity more often than not. Indeed, prior work in usable privacy and security suggests that many security features remain unused because stringent security measures are often antagonistic towards the specific goal of the end user at any given moment [92]. For example, while a user might want to quickly check her e-mail in between meetings, two-factor authentication *prevents* her from checking her e-mail immediately. Thus, people often reject security features when they expect or experience them to be weighty [4]. Consequently, typically only people who are especially dedicated to protecting their information use interruptive security features, and we know from Gaw et al.'s prior work that non-experts may perceive these early adopters as “paranoid” [45]. More formally, because early adopters of security features are likely to be perceived by others as behaviorally *different* (e.g., either paranoid, or in possession of expert knowledge), non-experts may perceive an illusory correlation [20], or an exaggerated relationship, between security feature usage and this behavioral difference. In turn, as non-experts consider themselves different from those who use security features, they may reject the use of security features. Moreover, this illusory correlation should only *strengthen* as more of these security-enthusiast early adopters use the feature because of the “availability heuristic”—a mental shortcut that biases people's judgments towards what is more frequently recalled [105].

The upshot of all of this is that the subjective perceptions of a security feature that propagates through social channels may work *against* its adoption, at least until enough of a potential adopter's more behaviorally similar friends start using the feature so that its use becomes normative.

What are the Gaps In Our Understanding? What are the Opportunities?

While there is a host of rich prior work in social psychology highlighting the importance of social influence in driving human behavior, and a host of rich prior work in usable security highlighting the reasons why end-users avoid using security tools, the background literature explicitly exploring how social processes affect privacy and security decisions remains surprisingly thin.

Indeed, to my knowledge, little work has looked at how social influence affects security sensitivity or how laypeople generally communicate about security and privacy (outside of Rader et al.'s study on security storytelling [81]). Yet, understanding how social influence affects security related behavior change and communication could improve our understanding of why security sensitivity remains low and could even help inform the design of social interventions that raise security sensitivity.

Similarly, while much background work *alludes* to the potential efficacy of social proof in heightening security sensitivity, there is a lack of work testing this potential. Part of the problem is that security behaviors have historically been kept secret to preserve the privacy of individuals. Still, as social channels are the primary way through which innovations spread [86], the hiding of social meta-data surrounding security behaviors has undoubtedly inhibited both the widespread adoption of security behaviors *and* research in studying social cues as a way to heighten security sensitivity. The little empirical data we *do* have about the effects of social influence on security related behavior change comes from work that only treated the social dimension in passing. Egelman and colleagues [40] included a social condition in their study on the effects of various types of password meters on convincing people to create stronger passwords. They found that a “peer pressure” password meter that showed participants how strong their passwords were relative to other users performed no better in increasing the strength of participants’ composed passwords, as compared to a standard password meter that told participants whether their passwords were “weak”, “medium” or “strong”. However, Egelman and colleagues’ “peer pressure” password meter measured participants’ passwords relative to strangers’ passwords for a completely different service, and provided little feedback as to whether a given meter reading was important enough to act upon (is it good or bad that my password is better than 50% of “others”?). In addition, their social intervention could only have an affect on participants’ motivation—not awareness or knowledge. Thus, *there is a strong need for more empirical work to build an initial theory of how social processes affect security behaviors.*

Finally, few security tools have been designed to be “social”—i.e., with an understanding of the social consequences of security tool use or with the intention of leveraging social processes to maximize their adoption. For example, even though people frequently share passwords and PINs [54,99], few security tools have been developed to be *inclusive*. Rather, most existing security solutions support sharing access through ad-hoc solutions, if at all: e.g., by having people create a “guest PIN” that they have to separately remember, or by sharing their original password or PIN which affords guests unrestricted access. Likewise, few security tools are built to be *observable*, so that their use can be seen by others (to maximize social spread) without compromising the original user’s security. Also, few tools allow people to act on their *sense of responsibility for their friends and loved ones* security. More generally, as Ackerman argues, there is a socio-technical gap between what is socially required of security and what has been technically feasible [1]. Given the explosive growth in sensing technologies, modeling techniques and interaction design, however, I believe that *there is a strong need and large opportunity for more systems design and development work that bridges this socio-technical gap between the social requirements of security and the technical features of security.*

In my dissertation, I will start with some mixed-methods empirical work exploring how social influence affects security behaviors and synthesize these findings into a set of descriptive and prescriptive insights. I will then design, implement and evaluate two socially inspired prescriptions—one nudging system and another authentication system—in order to understand how social design principles can be used to improve end-user facing cybersecurity systems.

Chapter 3: A Typology of How Social Influence Affects Security Behaviors

The contents of this chapter are drawn from a previously published paper: [The Effect of Social Influence on Security Sensitivity](#). Sauvik Das, Tiffany Hyun-Jin Kim, Laura Dabbish and Jason Hong. In Proceedings of the 2014 Symposium on Usable Privacy and Security (SOUPS'14) [28].

Summary

The first step in making security more social is to build a theory for social cybersecurity—or, to understand the relationship between social influence and security adoption, and to identify areas where current security tools can be improved. To build this understanding, I employed both in-depth qualitative and large-scale quantitative research methods. In this chapter, I present some formative work in which I used semi-structured interviews to ask people of various ages and backgrounds about their recent security-related behavior changes and communications [1]. **I found that social factors were key drivers of security-related behavior change**, accounting for nearly half of all reported behavior changes (e.g., using a PIN on one’s phone or enabling a Facebook security tool). Specifically, I uncovered five different social triggers that drive security-related behavior change: observing friends, a friend experiencing a security breach, social sensemaking, pranks / demonstrations of insecurity, and sharing accounts / devices. **The most prevalent social trigger for was observing friends**—i.e., people often started using security tools after observing friends and/or strangers use those same tools. Unfortunately, few security tools are built for this form of passive observability, and are thus unable to spread in this powerful and social way.

Motivation

Much prior work in usable privacy and security has looked at improving all parts of the security sensitivity stack—for example, through games for security education [98], browser extensions to make users more aware of phishing [114], more effective user interfaces for security tools [68], and faster or simpler ways to authenticate users [102]. Security sensitivity, nevertheless, remains low.

I argue that part of the problem is that we do not yet understand the *social* processes underlying people’s decisions to communicate about security and adopt security tools. In other words, security behaviors—as any human behavior—should be viewed within the context of a social system. Indeed, the social psychology and sociology literature illustrates that social influence, or our ability to affect other people’s perceptions and behaviors with our words and actions [21], plays a central role in how people behave—even specifically in changing their behavior or adopting a new technology or idea [21,88]. Rogers’ highly influential *diffusion of innovations* work, for example, has shown that social influence drives technology adoption [88]. Social processes, thus, should undoubtedly affect a user’s decision to follow security advice or adopt a security tool.

Nevertheless, the effect of social influence on decisions people make about security and privacy remains relatively understudied. Indeed, we do not yet know *how* social influence affects behavior change with regards to security and privacy. Understanding how social influence affects security related behavior change and communication should improve our understanding of why security sensitivity remains low, and, in turn, may help inform the design of social interventions and tools that can raise security sensitivity. To that end, I conducted a retrospective interview study aimed at investigating the following research questions:

***RQ1:** What role does social influence play in an individual’s decisions to use, discontinue use, and explore security tools and privacy settings?*

***RQ2:** Under what circumstances do people communicate about security and privacy?*

However, as these two research questions are fairly distinctive and inform distinctive directions for future work, in this chapter, I will focus on our findings with respect to **RQ1**. In Chapter 4, I will more deeply explore our interview findings with respect to **RQ2**.

Methodology

Semi-Structured Interviews

We constructed an IRB approved semi-structured interview protocol to probe participants about recent security related behavior changes. We elected a semi-structured approach so that we could concretize the discussion by directing participants' memories towards *changes* in behavior, while still allowing participants the flexibility to expand on their undoubtedly unique experiences. Our interview protocol probed participants about recent changes in (1) *mobile authentication*, or whether and why participants enabled, disabled, or changed authentication on their smartphones (e.g., from PIN to Password); (2) *application installation and uninstallation*, or whether and why participants decided to uninstall or halt installing applications because of privacy and security concerns; and, (3) *online privacy settings in social media*, or whether and why participants changed their privacy settings on the social media platform they most commonly used. We chose to explore three categories to uncover general trends across different types of security tools, and we chose these three categories specifically because they represented a broad range of behaviors representative of common security and privacy decisions made by most people on a fairly regular basis.

If participants reported a *specific* security-related behavior change, we asked them to explain further how the change was catalyzed—specifically, to discern between *social* and *non-social* catalysts for behavior change. Either way, we asked participants to explain, in detail, the context surrounding their decision to enact the change: Was the change brought about by a personal negative experience, or because of an article they read online? If they heard about a security incident through a friend, how did the friend broach the conversation? And, if a social process drove the change, we asked participants to clarify how the social process manifested—for example, did they seek out advice, or did a friend offer them unsolicited advice? We also asked participants whether and why they did or did not share their concerns, advice, or behavior change with anyone else.

We iteratively refined our protocol by piloting it with 5 people. All interviewers participated in the pilots in order to mitigate variation in delivery across interviewers and interview sessions. Questions that participants could not easily answer (e.g., hypotheticals) were culled through these iterations. Ultimately, our interview lasted approximately 45 minutes, and interviewees were compensated \$10 to participate.

Recruitment

We recruited participants from CBDR¹, an online recruitment tool that pairs research participants from Pittsburgh with research projects of interest. Participants were required to own a smartphone running Android or iOS, be an active user of any social media service, and be at least 18 years old. We went through three rounds of recruitment to recruit a variety of occupations and ages across our sample. For example, in our first round of recruitment, we predominantly interviewed students in their mid-twenties. Thus, in subsequent recruitment rounds, we specifically recruited older non-students. We stopped recruiting additional participants once we believed we had sufficient diversity in occupation, age, and security proficiency to capture a large cross-section of experiences with security-related behavior change and communication. In our case, we appeared to reach this point after interviewing 19 participants—indeed, after the first 15, every additional participant echoed experiences very similar to those previously reported by others.

Our participants ranged in age from 20 to 54 years old ($m=28.5$, $sd=10$). Seven out of the 19 participants were female. Furthermore, as we tried to recruit participants from diverse backgrounds, 10 of our participants were non-students from many different professional backgrounds. All participants used an Android ($n=12$) or iOS ($n=7$) smartphone and were frequent Facebook users. Fifteen of the 19 participants reported using Facebook daily, while the remaining 4 reported that they checked Facebook at least a few times every week. shows an overview of participants. Table 1 shows an overview of all 19 participants.

¹ <http://cbdr.cmu.edu/index.asp>

	Age	Gender	Race	Occupation	Phone OS	Phone Auth.	Social Media Usage
P1	28	Male	African American	Customer Service	Android	None	Daily
P2	22	Female	Asian	Unemployed	iOS	None	Daily
P3	22	Female	African American	Student	iOS	PIN	Daily
P4	22	Male	African American	Student	Android	None	Daily
P5	27	Female	Asian	Unemployed	iOS	None	Daily
P6	29	Male	White	Software Developer	iOS	None	Daily
P7	54	Female	White	Administrative Assistant	iOS	PIN	Weekly
P8	31	Male	Indian	Unemployed	Android	None	Weekly
P9	30	Male	White	Software Developer	Android	None	Weekly
P10	37	Male	White	Graphic Designer	Android	9-dot	Daily
P11	54	Male	African American	Chef	Android	None	Weekly
P12	20	Female	African American	Student	iOS	None	Daily
P13	24	Female	Indian	Graduate Student	Android	None	Daily
P14	25	Male	Indian	Graduate Student	Android	PIN	Daily
P15	21	Male	Indian	Graduate Student	Android	9-dot	Daily
P16	22	Male	Indian	Graduate Student	Android	9-dot	Daily
P17	34	Female	Asian	Unemployed	iOS	None	Daily
P18	20	Male	African American	Student	Android	9-dot	Daily
P19	20	Male	White	Student	Android	9-dot	Daily

Table 1. Interview participant demographics, occupations, use of authentication on their mobile phones, as well as social media usage.

Data Coding and Analysis

We recorded and transcribed, with consent, each interview, and used a qualitative data analysis program called Dedoose² to analyze the anonymized transcripts. We partitioned each transcript into a set of excerpts comprising of all instances of an *action taken*, a *decision made*, or, more generally, a *behavior changed* related to security or privacy. A representative example of behavior changes is P18's decision to rub-off the smudges on his Android device after a friend demonstrated that the smudges on his screen makes it easy for others to "crack" his Android 9-dot pattern:

"What I've been doing, I believe, after that scare with the nine dot, pretty much every time I turn off my phone, I put it in the pocket, I just kind of rub, just rub the smears off so you can't really see what direction I was going." (P18)

² <http://www.dedoose.com>

The second set of excerpts was a collection of all *specific* instances of communication about security and privacy, which we will refer to as the *communications*. An example excerpt comes from P14. After he received spam mail from a friend’s e-mail account, he mentioned:

“I told my friend that this is something weird that came from your account. This is not what you would be probably into.” (P14)

In total, from our 19 transcripts, we extracted $n=114$ behavior change excerpts. Excerpts were usually just answers to pointed questions, but to ensure robustness, two of the research group mutually agreed on all partition points for each excerpt.

We used these excerpts as our units of analysis—though, occasionally, we aggregated data across participants where it made sense (e.g., in determining how many participants actually changed their behavior as a result of a social process). We used an iterative, open coding process [75] to code the data, constructing codes where patterns naturally emerged and refining the codes iteratively until we reached consensus. Our goal was to understand the effect of social influence in driving *behavior changes*—which, in turn, means understanding the effect of social influence in modulating *security sensitivity*.

Concretely, two researchers independently and openly coded a random subset of 20% of the excerpts. These openly generated codes were collaboratively synthesized into a set of high-level codes that three of the research team then used to code the remaining excerpts. Upon completion, the coding team discussed potential extensions to the coding scheme that arose from coding the new examples. If a change to the scheme was made, the coding team re-coded the full set of excerpts with the new scheme. We required two coding iterations to come to consensus.

From the 20% overlap of excerpts, overall inter-coder agreement was 85% (calculated as the number of overlapping excerpts where codes matched divided by the total number of overlapping excerpts). In cases of discrepancies, the coders discussed the discrepancies until agreement was reached, following standard practice. Inter-coder agreement for *each* applied code can be found in Table 2, and all exceeded the 0.7 threshold commonly held to be acceptable in qualitative research [75].

<i>Code</i>	<i>Inter-Coder Agreement</i>
<i>Behavior Change: Social or Non-Social</i>	0.93
<i>Behavior Change: Trigger Event</i>	0.87
<i>Behavior Change: Raised Awareness</i>	0.87
<i>Behavior Change: Raised Motivation</i>	0.80
<i>Behavior Change: Raised Knowledge</i>	0.80

Table 2. Inter-coder rating agreement on 20% of the behavior change excerpts.

Results

First, we wanted to know if social processes often drove security related behavior changes, so we coded each *behavior change* excerpt as being driven by a *social* or *non-social* process. Excerpts were coded as being driven by a social process when the reason for the behavior change was social, and, importantly, if the social process was *clearly* reported by the participant in the transcript. For example, when asked about why he first enabled a PIN on his iPhone, P6 stated:

“When I first had a smartphone I didn’t have a code, but then I started using one because everyone around me I guess had a code so I kind of felt a group pressure to also use a code.” (P6)

As the underlying reason for the behavior change was a social process (observing one’s friends) and was stated as such, we coded that behavior change as social. An example of a non-social behavior change comes, again, from P6. When asked why he changed his Twitter password, P6 responded:

“Diversification of passwords. I had the same password for every service so I wanted to pick a stronger password for... the service, yeah.” (P6)

While P6 *could* have learned about the need for password diversification from friends, as he did not explicitly confirm this speculation, we coded the excerpt as non-social.

In all, out of the 114 behavior change excerpts, we coded a substantial 48 as being explicitly driven by some form of social influence. Furthermore, most participants (17 out of 19) reported at least one action taken, decision made, or behavior changed that was driven by social influence. Of note, however, is that the 48 examples of socially driven behavior change did not come uniformly from all of our participants. Notably P2 and P10 reported the largest number of socially driven changes at eight, each. It is important to keep this bias in mind in any quantitative interpretation of our findings.

In all, these results suggest that social influence already plays a strong role in driving security and privacy related behavior change—even without any explicit social interventions. Next, we wanted to understand when and how social influence is effective at driving these behavior changes.

Social Triggers in Driving Behavior Change

To explore *when* social influence drove behavior change, we open coded the *triggers* for behavior change excerpts coded as “social”. We found five primary social triggers for behavior change: *observing friends, social sensemaking, pranks and demonstrations, experiencing security breaches, and sharing access*. Table 3 lists all triggers, their frequency and their description.

Next, to answer *how* social processes enacted behavior change, we also coded whether or not the socially driven behavior change examples in our dataset affected any part of the security sensitivity stack. Specifically, we asked the following:

<i>Trigger</i>	<i>N</i>	<i>Description</i>	<i>Example</i>
<i>Observed friends</i>	14	Observing people around them engaging in a particular security behavior and emulated those people.	“So when I was an undergrad I’ve been using it since then. And this four digit everybody started using it and it was a hype. And we had it.” (P14)
<i>Social sensemaking</i>	9	Discussing concerns with friends/loved ones to determine the right behavior.	“I mean, like, one of my friends told me that you could alter the privacy settings so that, like, not everyone can look up your profile and not everyone can, like, try sending messages to you.” (P15)
<i>Prank/ Demonstration</i>	8	Friends/loved ones hacked into his/her account, demonstrating they were insecure.	“Yeah, like my laptop was in my room. I walked out of my room and someone walked by and saw my Facebook and thought it would be funny to put something up.” (P19)
<i>Security breach</i>	6	Someone hacked into his/her account or information was shared too widely.	“I did change that within the past week. The girlfriend was reading all of my mail, which is also a privacy concern” (P10)
<i>Sharing access</i>	3	Sharing access to a device or account with another person leading to need for better security.	“There are sometimes when you have to tell your friends what is my PIN number because they are a very good friend of yours and they have to make a call and I can’t go every time and just unlock this for them.” (P14)

Table 3. Social triggers for behavior change derived from our iterative open coding process.

Raised Awareness: Did the social process raise the participant’s awareness of a new threat and/or security tool?

Raised Motivation: Did the social process raise the participant’s motivation to protect him or herself against a security threat?

Raised Knowledge: Did the social process raise the participant’s knowledge of how to use a security tool or method?

Importantly, we only answered “yes” to those questions if the *social* process mentioned in the excerpt was the reason for the heightened security sensitivity. For example, P16 mentioned that his Facebook account getting “hacked” resulted in him changing many of his passwords every 6 months at the advice of his friends, who he sought out for advice after the incident. In this example, the social process of P16 speaking with his friends raised his *knowledge* but not his *awareness* or *motivation*. It was the non-social process of experiencing a breach that raised his awareness and motivation.

For most (44 of 48) reported examples of socially driven behavior change, we found that the social process triggering the behavior change did, in fact, raise *some* form of security sensitivity. In fact, many examples raised *all* points of the security sensitivity stack. For example, P18 recalled advice he received on password composition after asking his friend to share a password:

“When I was working this summer, one of my co-workers told me about the whole algorithm thing. One, it just helps you I guess have different passwords. It helps you recall them easier based on I guess the type of profile. I guess you can cater, you can change your algorithm, depending on I guess what you want to be in it. But ever since I started using it.” (P18)

In this example, the social process of P18 asking his friend about how to compose a password increased his awareness of a new method of password composition, his motivation to update his own method of password composition, and his knowledge of how to improve his method of password composition. In the text to follow, we describe each social trigger we found in our data for security related behavior change. Furthermore, as a descriptive aid, we plotted how frequently different social triggers raised the different components of security sensitivity in Figure 3.

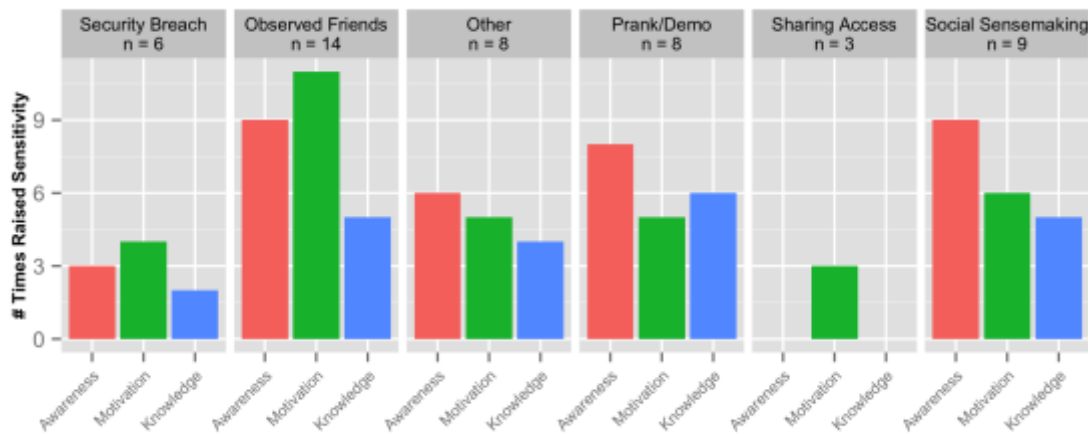


Figure 3. The number of times each social trigger for behavior change reported by our sample raised any of the three parts of the security sensitivity stack: awareness, motivation, or knowledge.

Observing friends (14/48 examples)

Most frequently, our participants reported changing their behavior after observing the actions of friends or others around them. In other words, participants changed their behavior after finding *social proof*—or, cues on how to act based on the actions of others [21]. For example, one participant in our sample adopted the 9-dot authentication method on his Android phone because his friends also used it. Additionally, as previously illustrated, P6 adopted a PIN because he felt “group pressure” to do so after observing everyone around him use authentication. This finding appears to be well supported by the background literature on technology adoption, which lists *observability* as a key criteria for an innovation to spread rapidly through social channels [88].

In certain cases, other forms of social influence apart from social proof appeared to be at play—specifically the social influence concepts of *liking*, or our tendency to follow the advice of those we like and those like us, and *authority*, or our tendency to follow the advice of those we consider to be authority figures [21]. For example, one participant indicated that she adopted a PIN code for her iPhone wholly because her mother, who she considered technically savvy, also had a PIN:

“My mother has-- she had an iPhone before I did, so she always had the block on hers, so I just kind of the... I think just because I saw her doing it, so it kind of just felt like it was something I had to do too.” (P3)

Observation influenced behavior change for mobile authentication more often than the other specific topics we asked about in our interviews, probably because it is relatively easy to observe others authenticating onto their phones compared to observing others update their social media privacy settings or uninstall an app.

Looking at Figure 3, participants who observed others use security tools often were themselves *motivated* to start using those tools (11/14 examples). Furthermore, participants often became more *aware* of security tools after observing others’ using those tools (9/14), but only occasionally gained *knowledge* of how to use the observed tools and methods by observing others (5/14).

Social Sensemaking (9/48 examples)

The second most frequent social trigger reported by our sample was *social sensemaking*—or, the process of making sense of a security system, tool, or threat by discussing concerns with others. We termed these triggers social sensemaking because they were similar in form and purpose to discussions, observed by Weick et al., among members of an organization who attempted to resolve uncertainty about recent novel events in their environment [110].

Participants often reported having discussions to resolve ambiguity in news and hearsay about security. The aim of these discussions was usually to find the correct or appropriate way to act to achieve the desired level of privacy or security within a system or with a security tool. In many cases, these discussions were prompted by a sudden infusion of uncertainty—for example, news articles about a novel security threat or gossip about anomalous security breaches others had experienced. Participants discussed these novel threats with others to share information about the threat, assess its veracity, and determine whether and how to change their behavior in response. For example, one participant in our dataset reported becoming more restrictive with posting to Facebook in response to a sudden, alarming, but unclear threat of all timeline posts becoming public:

“So yeah. I recently, like, a day or two, day before yesterday, I went through an ordeal. I don’t know if it’s fake or it’s real, but somebody mentioned that all his private messages, they became public. Like, his messages with a friend. And it was like he had never thought of putting it on wall. And it suddenly opened his Facebook and everything was on his...I don’t know if it’s a real thing. And somebody mentioned in a comment that it happened with him as well, few days back.” (P16)

P16's example is another illustration of social proof based social influence affecting an individual's security behavior: facing an ambiguous threat, P16 observed his friends for cues on how to act.

Social sensemaking also occurred when a participant wanted to understand a particular function within a system—for example, Facebook privacy settings. This need for specific information resulted in discussion and information sharing that exposed novel functionality or methods for protecting oneself against threats—often increasing participants' *knowledge* about the system (5/9 examples) and eventually leading to behavior change as a result. For example, one participant updated his privacy settings after a discussion that revealed novel system functionality:

"I mean, like, one of my friends told me that you could alter the privacy settings so that, like, not everyone can look up your profile and not everyone can, like, try sending messages to you. As in you can go to the privacy settings tab. And then, you could actually change it. Because I didn't know that you could do it, before. I mean, I just thought that it was default that everyone could look at your profile." (P15)

Social sensemaking also made participants more aware of available security tools (9/9), and the discussions would frequently motivate participants to act on their newly acquired knowledge (6/9).

Prank/Demonstration (8/48 examples)

The third most prevalent social trigger for the behavior changes reported by our participants was pranks and demonstrations—i.e., friends or loved ones cracking participant's accounts and devices as a prank, or to demonstrate that they were being insecure. Often, these pranks were explicit demonstrations to prove to the victim that their current security strategy or behavior was insecure. For example, one participant in our sample described a co-worker breaking into his phone to show the vulnerabilities of 9-dot authentication:

"One of my, when I was interning, engineering company, one of my friends and a fellow intern came to my desk, just unlocked my phone. I was surprised. I was like, "Hey, how'd you do it?" He put it against the sunlight and he saw I guess the smudges my finger left. He just followed the direction. Yeah, he had access to my phone." (P18)

Other prank examples reported were simply driven by opportunity—for example, a friend gaining unauthorized access to the participant's account because they left their Facebook account open on an unprotected device. Indeed, several of our participants were motivated to change their security behavior after their friends accessed their social media accounts and posted embarrassing information on their behalf. For example, one participant experienced this type of prank after leaving his laptop open and unprotected in his dorm room:

"Besides just my friends getting into my phone or on my Facebook and that's more from just me leaving my Facebook open or something if I walk out of the room and they just put up a funny status or something like or even just look through my messages or something like that. But nothing too threatening, more like practical joking side of it. But once that happens, I usually change my password immediately as would all of my other friends would too." (P19)

Pranks appeared to be quite effective at raising participants' security sensitivity. In all cases (8/8 examples), participants were made aware of a security threat and, in most cases, participants were instantly motivated (6/8) to update their behaviors to prevent a reoccurrence of the prank. Pranks aimed at demonstrating insecure behavior were also effective at raising participants' knowledge (5/8), as they were often followed up with direct or indirect lessons to prevent the breach from reoccurring—for example, the screen smudge "hack" reported by P18 taught him to wipe out the smudges from his phone screen periodically.

Experienced a security breach (6/48 examples)

Another prominent social trigger was experiencing a security breach—when participants or someone they knew had an account or device accessed by a stranger, or otherwise had information shared with unintended parties. In these examples, the victims of a security breach solicited advice from friends and loved ones, simultaneously spreading *awareness* (3/6 examples) of a new security threat, and *motivating* (4/6) behavior changes by grounding it in a real example of harm.

One participant initiated a new practice of updating his password on a monthly basis following his Facebook account getting breached, because his friend recommended that course of action:

“Because once I got my account hacked. And I was [doing my] bachelor’s in a city, so yeah. After that I was more precautious regarding the same. And I’ll keep changing my password, so on a monthly basis [because] My friends, actually they recommended me to do so. Like there’s one of my friends used to do it. He said it’s better to be safe than sorry, so...” (P16)

Sharing access (3/48 examples)

Another general social trigger was behavior change triggered by sharing a device or account with a friend or loved one—for example, modifying a password after allowing a friend to check their phone. These changes were a reflexive response to the fact that what participants desired to generally be private was now more widely available because of a transient need to share access. For example, one participant let her son use her phone and updated the passcode afterwards:

“One of my boys wanted to use my phone for something so I gave them my passcode. And not that I have anything that I don’t care for them to see or anything, but after they did that then I changed it again because I just didn’t want anybody to just-- I don’t care if it’s them or not. I don’t want them to just be able to pick up my phone and do what they want with it.” (P7)

While these triggers rarely raised awareness (0/3 examples) or knowledge (0/3), they seemed to be motivate participants to make a change (3/3).

Other triggers (8/48 examples)

Eight other instances of behavior change reported by our sample were triggered by other experiences, usually conversations or recommendations—for example, an authority figure recommending the use of authentication, as mentioned by P8 when asked why he first enabled mobile authentication:

“I think my boss at the time had it and he recommended it, because he leaves his phone at his desk.” (P8)

Likewise, P10 mentioned adopting anti-virus software after receiving a recommendation from a friend who he considered a security expert, and P13 mentioned that she stopped using Google Chrome for financial transactions because two of her security expert friends informed her that the version of Chrome she used insecurely stored information. These recommendations often raised participants’ awareness of, motivation to use and knowledge of how to use a new security tool.

Importantly, however, recommendations from authority figures didn’t always result in behavior change. P13, for example, mentions that she ignored her boss’s advice to have different passwords for different accounts because it would be hard to remember all those passwords. Nevertheless, the advice did raise her awareness of proper security practices.

P7 reported re-activating the PIN for her iPhone because a family member asked her why she deactivated it in the first place, urging her to reconsider. The conversation didn’t raise her awareness or knowledge, but re-upped her motivation to use a security tool with a bit of social proof.

Interestingly, another participant mentioned installing anti-virus software on her laptop simply because she felt guilty, after conversing with others who attended her university's cybersecurity awareness fair, for not using software that her school provided:

"I also felt guilty that I have all this free stuff I could install to protect my computer, and all this stuff I could do that's smart and I wasn't taking it." (P12)

The guilt inspired behavior change reported by P12 is emblematic of the reciprocity principle of social influence, which suggests that people are more likely to follow the suggestions of those who did them a favor—even an unsolicited one [21].

Importantly, one participant reported how a social process urged her against behavior change (but was still responsible for a decision she made about security). P17 mentioned that she did not follow her security-expert husband's advice to delete unused and obscure online accounts because she noticed that her friends, who did not follow the advice, never experienced a security breach:

"I don't think it will be dangerous. Maybe I didn't see this kind of news or my friend didn't get some trouble when they didn't set password. Like, my friends sometimes they usually have a lot of different accounts, the same as me. But they didn't get any trouble. So I think maybe it will not be dangerous." (P17)

In this way, P17's friends' lack of a security breach offered her social proof that it's okay to ignore her husband's security advice.

Discussion

In summary, I interviewed 19 participants about specific, recent security and privacy related behaviors they had changed, actions they had taken or decisions that they had made. From these interviews, I extracted and analyzed 114 examples of behaviors changed, actions taken, or decisions made related to security and privacy. My results introduce a typology of social interaction around cybersecurity behavior. First, I confirmed that social processes are an important influence on cybersecurity behavior change—indeed, a large number of behavior changes (48 / 114) reported by my sample were driven at least partially through social processes. Specifically, I identified five common social triggers for security related behavior change—observing and learning from friends, social sensemaking (discussing ambiguous security threats with friends to determine the relevance of the threat and a clear course of action), pranks and demonstrations, experiencing a security breach and sharing access to a device with others. All of these social triggers appeared to heighten security sensitivity in some way—either by increasing participants' awareness of a new threat or security tool, motivating participants to protect themselves, or increasing participants' knowledge of how to protect themselves. Taken together, these findings lend some support to the notion that social influence, especially in the form of *social proof*, *authority*, *liking*, and *reciprocity*, can be potent in raising security sensitivity—a result that supports the allusions of prior work [45,81,99].

Opportunities

These results also highlight some opportunities to leverage social processes to drive security-related behavior change:

Creating teachable moments out of negative experiences. My results emphasize the influential nature of a specific negative experience in raising the security sensitivity and, in turn, changing the cybersecurity behavior of victims and those around them. Interestingly, friends and loved ones appeared to at least indirectly take advantage of this fact, often breaking into others' accounts to prove to that person that s/he was not fully protected. This notion of pranking by friends and family can also be considered as an effective way to create a *teachable moment*: an instance during which an intervention to teach people about better security behaviors is especially likely to be effective [71]. In other cases, pranks were not necessarily meant to directly educate victims, but were used as a form of *hazing*. Either way, the breach elicited a similar reaction—both the victims of these pranks and the people around them with whom they shared the experience became more aware of and motivated to address their own security vulnerabilities.

Creating more observable security tools. The observability of security features and methods also proved to be important in driving behavior changes through social processes. Indeed, observing friends was the most frequent social trigger for behavior change. Nevertheless, most security features and methods are inherently unobservable—for example, password composition methods. When P18 learned of a new way to compose passwords from his friend, he immediately started utilizing this new composition policy. However, only two of my participants mentioned talking about password composition policies, suggesting that there is much room for improvement in leveraging observability to raise security sensitivity.

However, simply increasing the observability of all security features may not be the best solution. First, security settings have historically been private—and for good reason. Indeed, past work by Gaw et al. [45] found that people who encrypted e-mail were often considered paranoid unless they were in a role where they handled sensitive company data, suggesting an illusory correlation [20] between security feature usage and paranoia. Indeed, as early adopters of security features are likely those who are especially concerned about their security—and, thus, are the most likely to be considered as paranoid by lay users—it is possible that making security decisions and behaviors perfectly observable might work *against* security sensitivity. First, potential adopters may look at early adopters and find tenuous social proof that only “paranoid” people use a security feature. Second, I also saw evidence that social processes can work against a user following advice if it seems like none of their friends are affected by a threat—for example, it is possible that when a useful security behavior has few existing adopters, others might see the absence of adoption as social proof against the behavior.

To best leverage observability, therefore, it seems that we should create security tools that are more visual and amenable to conversation, such that non-experts can passively raise their *awareness* and *motivation* by observing their friends, and then raise their *knowledge* by asking about security.

Creating security tools that facilitate sharing access with others. Finally, several participants mentioned sharing access to accounts and devices as a prompt to change their authentication secrets. While the result of updating one’s password after sharing one’s device with others is desirable, this practice suggests a broader weakness of many present security tools—the assumption that people would never want to share their accounts and devices with others temporarily or regularly. As pointed out by prior work on home data sharing [74] and password usage in daily life [54], as well as illustrated by my own interviews, this assumption of non-sharing is flawed. Thus, in addition to making security tools more *observable* it seems that we should also make these tools more *socially inclusive* to better support these sharing practices. In turn, by making security more inclusive, we can also combat the perception of “paranoia” associated with good security behaviors [28,45].

Chapter 4: Understanding How Security Information Is Communicated

The contents of this chapter are drawn from a previously published paper: [The Effect of Social Influence on Security Sensitivity](#). Sauvik Das, Tiffany Hyun-Jin Kim, Laura Dabbish and Jason Hong. In Proceedings of the 2014 Symposium on Usable Privacy and Security (SOUPS'14) [28].

Summary

I showed that social influence is a key driver of people's security and privacy behaviors. A number of these social triggers for behavior change, however, required active communication between multiple people about security tools and/or threats. To get a better understanding of these communications, I next investigated whether and how people communicate about security. In this chapter, I report on results from the second half of the same interview study in which I investigated communications about security and privacy people have with their friends, family and other social relations. I learned that communications about security are scarce, with many participants, including several security experts, reporting that they did not talk much about security at the risk of being boring or sounding preachy. From the conversations that participants *did* report having, however, I construct a typology of security communications. Specifically, my data suggests that conversations about security occur primarily to warn others of security threats or to teach others about how to protect themselves from a threat. Thus, it seems that **people feel accountable for the security of their loved ones**.

Motivation

In the previous chapter, I showed that social influence is a powerful tool to affect security and privacy related behavior change. But, it remains unclear: How prevalent is socially driven security-related behavior change? Socially driven change is the result of an interaction between two or more individuals—but those interactions are rare in the domain of security and privacy. Indeed, when asked why he didn't share his concerns about the U.S. government's pervasive surveillance (NSA PRISM) program, one of our participants stated: *"That's one thing I will never talk about."* Similarly, when asked about whether he has warned friends about a malicious smartphone application he uninstalled, another stated: *"Especially online. In person, it depends on the context. It does become a boring subject."* The realization that conversations about security remain rare—and, thus, so too does the potential for socially driven behavior change related to security—begged the question: Under what circumstances do conversations about cybersecurity occur? In this chapter, I explore the second research question we asked in our interview study:

RQ2: Under what circumstances do people communicate about security and privacy?

Methodology

In the same semi-structured interview in which we asked participants about specific security-related behaviors they undertook and the reasons they made those changes, we also asked participants if they could recall *specific* conversations they had about security and privacy. For example, if a participant mentioned that a conversation she had was the reason she started using a PIN, we further probed that participant to provide more details about the conversation—e.g., who told her to use a PIN? How did that conversation start? In addition, to capture security-related conversations that did not fit into our pre-constructed themes of mobile authentication, app installation, and social media privacy settings, we also asked participants more open-ended questions about conversations related to security and privacy. Did they ever share information about security or privacy? If so, what did they share, with whom, and why? These more general questions were asked at the end of the interview.

By focusing on specific conversations about security and privacy (e.g., "I told my mother to update her privacy settings"), rather than general conversations (e.g., "People usually tell me to update my password"), we were often able to uncover the specific context of a conversation (e.g., a catalyst and goal for the conversation).

To reiterate, we recruited 19 participants to interview from the greater Pittsburgh area with CBDR and iteratively refined our interview protocol with 5 pilot participants before conducting the actual interview. Our participants ranged in age from 20 to 54 years old ($m=28.5$, $sd=10$), and seven were

female. For more specific details about participant recruitment, demographics, and compensation, please refer to the Methodology section of Chapter 3: A Typology of How Social Influence Affects Security Behaviors, as these insights are obtained from the same interview study as that of Chapter 3.

Data Coding and Analysis

To reiterate the previous chapter, we recorded and transcribed, with consent, each interview, and used a qualitative data analysis program called Dedoose to analyze the anonymized transcripts. However, for this analysis, we identified excerpts pertaining to *specific* instances of communication about security and privacy. An example excerpt comes from P14—after he received spam mail from a friend’s e-mail account, he mentioned:

“I told my friend that this is something weird that came from your account. This is not what you would be probably into.” (P14)

We classified this excerpt as a communication excerpt because the participant explicitly mentioned conversing with a friend about something that could have implications for privacy or security. In total, from our 19 transcripts, we extracted $n=118$ communication excerpts. Excerpts were usually just answers to pointed questions, but to ensure robustness, two of the research group mutually agreed on all partition points for each excerpt.

We used these excerpts as our units of analysis—though, occasionally, we aggregated data across participants where it made sense (e.g., in determining how many participants reported having a certain type of conversation). We used an iterative, open coding process [75] to code the data, constructing codes where patterns naturally emerged and refining the codes iteratively until we reached consensus. Ultimately, our goal during the coding process was to better understand the triggers and reasons underlying communications about security and privacy.

Concretely, two researchers independently and openly coded a random subset of 20% of the communications excerpts. These openly generated codes were collaboratively synthesized into a set of high-level codes that three of the research team then used to code the remaining excerpts. Upon completion, the coding team discussed potential extensions to the coding scheme that arose from coding the new examples. If a change to the scheme was made, the coding team re-coded the full set of excerpts with the new scheme. We required two coding iterations to come to consensus.

From the 20% overlap of excerpts overall inter-coder agreement was 79% (calculated as the number of overlapping excerpts where codes matched divided by the total number of overlapping excerpts). In cases of discrepancies, the coders discussed the discrepancies until agreement was reached, following standard practice. Inter-coder agreement for *each* applied code can be found in Table 4. Inter-coder agreement of codes on a 20% random sample of communication excerpts., and both

<i>Code</i>	<i>Inter-Coder Agreement</i>
<i>Communication: Catalyst</i>	0.71
<i>Communication: Reason</i>	0.86

Table 4. Inter-coder agreement of codes on a 20% random sample of communication excerpts.

exceeded the 0.7 threshold commonly held to be acceptable in qualitative research [75].

Results

To understand the conditions under which conversations about security and privacy occur, we open coded excerpts about communication to surface triggering events for the interaction (*catalysts*) and the goal of the conversation (*conversation goal*).

Catalysts for Security Related Communication

We observed six primary catalysts for security related conversations, as summarized in Table 5. Below, we summarize each of these conversation catalysts in turn.

Catalyst	N	Description	Example
<i>Observed insecure or non-private behavior</i>	15	Noticed that someone was being insecure.	<i>"Right now I have ignored this storing passwords on my cell phone. He was like, 'Don't do this. It's dangerous.'" (P7)</i>
<i>Observed Novel Behavior</i>	11	Noticed a new security tool / method.	<i>"[I] see a lot of fancy password protection programs on [my co-workers] laptops. Like special files being encrypted. I'm like, 'What's going on?'" (P11)</i>
<i>Sense of obligation</i>	15	Shared information out of obligation to protect others.	<i>"When I was younger, I remember my parents always telling me, like I'm sure everyone's parents tell them, to be very careful about who they give their Social Security number to. So, that's always like in my head, like if someone asks me for that, I'm just like, uh, no." (P14)</i>
<i>Negative experiences</i>	33	Experienced a security or privacy breach	<i>"Yes, my data got stolen. My photo got stolen on Facebook. I spoke to a couple of my friends. The only thing I could do was report abuse." (P6)</i>
<i>Configuring settings</i>	14	Had to set up security for a new device, account or security tool.	<i>"He was asking about Facebook, and he's a businessperson, so social media is somewhat of a new thing to him, and I think Facebook was-- he was just curious about it and how he could use it to kind of help his business and stuff like that. So..." (P20)</i>
<i>News articles</i>	15	Read a news article.	<i>"Well, before, I did not even know like I need to pay attention to this. Like I was aware of this, but I just did not know it was such a big deal. Then later, like I saw a topic, like online articles talking about that _____, talking about that, and that's when I went to the setting of like Facebook to change some." (P5)</i>

Table 5. Conversation catalysts derived from our iterative open coding process.

Negative experiences (33/118)

Negative experiences were by far the most common catalyst for security conversations reported by our participants. These negative experiences could take many form, but often involved at least one of the conversation partners directly experiencing a security or privacy breach themselves. Indeed, many participants reported having conversations with friends and loved ones after experiencing a security breach. For example, one participant sought advice from friends after she received a friend request, on Facebook, from a fake profile using her own picture:

"Yes, my data got stolen. My photo got stolen on Facebook. I spoke to a couple of my friends. The only thing I could do was report abuse." (P6)

Observing insecure or non-private behaviors (15/118)

Often, participants reported starting a conversation in response to observing what they believed was non-secure behavior, such as a friend or family member oversharing on social media:

"One of the reasons we talked about it is because I saw so many people post things on Facebook. A lot of times it's unnecessary things, you know, like just what they did today, "Oh, I had an amazing day," or, "I had a great dinner," and I was just

talking to my husband, like why they-- I don't understand like they do that, like why they like to post things on Facebook to so-called to share.” (P5)

In this case, a participant, P5, observed her connections on Facebook engaging in what she considered “oversharing.” This observed behavior prompted her to initiate a conversation with her husband to understand why people engaged in this non-private behavior.

Sense of accountability (15/118)

A sense of social accountability also frequently prompted conversations about security. Curiously, this sense of accountability was not limited to any one social role. Rather, we observed that many different entities who held many different social roles all appeared to feel a sense of accountability for others’ security and privacy. For example, parents lectured their children about security and privacy best practices:

“When I was younger, I remember my parents always telling me, like I’m sure everyone’s parents tell them, to be very careful about who they give their Social Security number to. So, that’s always like in my head, like if someone asks me for that, I’m just like, uh, no.” (P14)

Likewise, managers informed their employees about how to manage company data because it was a part of their responsibilities. One participant described this type of interaction with his boss:

“When I was at work, I was given some sensitive documents, and I was told I couldn’t send them over e-mail. I had to use a flash drive to move them over, encrypt them, then send them in e-mail.” (P18)

Curiously, even large organizations, such as entire universities, appeared to feel a sense of accountability for its students. For example, one student talked about her university providing security solutions and advice in an annual security fair that she attended:

“They give us LoJack and all these different things you can get at the computer center. So we did talk about that. Like, locking up our computers and changing our passwords and stuff and being careful with the Wi-Fi.” (P12)

Reading news (15/118)

Unsurprisingly, news articles or other press about security and privacy breaches also frequently triggered conversations. For example, one participant read and subsequently shared an article on social media about how over sharing could lead to identity theft and, more darkly, black market organ trading:

“I know there’s like news talking about girls they are just so crazy about telling people on the social media where they are every minute, what they are doing every minute. So some criminals they actually use the information and just like kind of how do you say they found the girl according to her shared information online every minute. [...] So I shared this article just to let my friends see just don’t do it very often because I saw some of my friends on Facebook she did this really often like telling everybody what she was doing and what she had and where she was and like that.” (P2)

This link between online and offline crime can potentially make the consequences of poor computer security and online privacy practices more concrete—if extreme.

Configuring settings (14/118)

Another frequent conversation catalyst about security and privacy was the need to configure security and privacy settings on a new device, application or account. For example, one participant reported asking a friend for advice when a Facebook application asks for access to protected information:

“So there are many applications and Facebook would say that if you want to access them, there’s a pop-up saying, “Allow,” like, it will access all your information and stuff. So I asked him if I should go for it or not, and he tells me if it’s worth going. Like, “Is it reliable or not?” (P16)

In general, participants frequently started conversations when setting or re-setting Facebook privacy settings (P13, P14, P16). In addition, many participants reported parents or older friends initiating conversations when they were setting up new computers or social media profiles for the first time (P4, P10, P15).

Observing novel behavior (11/118)

Finally, people appeared to initiate conversations about security or privacy after observing novel behaviors—for example, a new, visually appealing authentication technique. Indeed, one participant was stopped in a coffee shop and asked about the 9-dot authentication on his Android phone:

“We were just sitting in a coffee shop and I wanted to show somebody something and [they said], “My phone does not have that,” and I was like, “I believe it probably does.” (P10)

In general, most security and privacy behaviors are designed to be invisible which prevents lay people from observing what experts do to protect their security. As a result, there are few vectors for lay people to converse with experts about their security habits and behaviors.

Conversation Goals

Conversations typically have agendas or goals. Thus, we next coded our 118 communication excerpts for conversation goals to better understand what people wanted to achieve from the conversation. Was it to warn others about potential threats, edify others about security tools or seek advice on how to configure security settings? During our open coding process, we identified seven distinct types of conversation goals, summarized in Table 6 below.

<i>Goal</i>	<i>N</i>	<i>Description</i>
<i>Notify or warn</i>	32	Notify or warn others of a potential security or privacy threat.
<i>Prank or Demonstrate</i>	5	Demonstrate insecure behavior by hacking into a friend’s account or device.
<i>Share solutions</i>	14	Share solutions, tools, and best practices (e.g., sharing how one composes his/her own password).
<i>Vent</i>	8	Seek social support / commiserate the experience.
<i>Offer advice</i>	19	Offer specific advice to others (e.g., update privacy settings, change password).
<i>Seek advice</i>	18	Ask for specific advice about security / privacy.
<i>Storytelling</i>	12	Topic was interesting/shocking/otherwise made for a good story.

Table 6. Conversation goals derived from our iterative open coding process.

Identifying a typology of communications: The Interaction of catalysts and goals
 Thus far, we have identified a set of conversation catalysts and goals. The interaction between these catalysts and goals affords us a typology of security and privacy communications. To construct this typology, we started by cross-tabulating catalysts and conversation goals. The results are shown below in Table 7:

	<i>Offer Advice</i>	<i>Share Solution</i>	<i>Vent</i>	<i>Seek advice</i>	<i>Notify or Warn</i>	<i>Storytelling</i>	<i>Prank or Demonstrate</i>	<i>Other</i>	Total
<i>Sense of Accountability</i>	8	2	0	0	2	3	0	0	15
<i>Observing insecure or non-private behavior</i>	4	0	1	0	7	0	2	1	15
<i>Negative experiences</i>	3	3	5	7	10	2	2	1	33
<i>Configurating settings</i>	2	2	1	8	0	0	0	1	14
<i>Reading news</i>	1	0	0	0	8	3	0	3	15
<i>Observing novel behavior</i>	0	5	0	3	0	2	0	1	11
<i>Other</i>	1	2	1	0	5	2	1	3	15
Total	19	14	8	18	32	12	5	10	118

Table 7. Co-frequency of catalysts for conversations about security and privacy (rows) and reasons for starting the conversation (columns).

For brevity, we focus here on the six most prevalent and interesting combinations, summarized in Table 8. These six combinations grouped into two broad categories of conversations, distinct in terms of their catalyst, focus and goal—*warnings* and *teachings*.

Name	N	Catalyst	Content
Warnings			
<i>Cautionary tales</i>	10	Negative experience	Notify / warn
<i>Targeted warnings</i>	7	Insecure behavior	Notify / warn
<i>Spreading the news</i>	8	News article	Notify / warn
Teachings			
<i>Lectures</i>	8	Sense of accountability	Offer advice
<i>Configuration help</i>	8	Configuration	Seek advice
<i>Social learning</i>	5	Novel behavior	Share solution

Table 8. The most frequent conversations about security and privacy, based on the catalyst and content.

Using this typology of security and privacy related conversations, we have enough context to answer our second research question: *under what circumstances do conversations about cybersecurity occur?*

Warnings

Warnings were meant to raise awareness of a specific, immediate threat that had come to the attention of the conversation initiator. These warnings took three forms, varying in their catalysts, but resulted in a notification about a novel threat: cautionary tales, targeted warnings, and spreading the news.

Cautionary tales (10/118 examples)

The most common catalyst-goal combination reported by our participants was what we called **cautionary tales**—a conversation triggered by a negative experience on the part of the conversation initiator or someone close to the initiator, with the goal of warning friends and loved ones about the threat. These conversations often involved sharing information about a recent security breach so that others could judge if their accounts or information were in any danger. In several cases the conversation was a response to an out-of-character behavior on the part of a friend or family member. For example, when asked about why he decided to reach out to his friend about a potential security breach, P11 mentioned:

“Because, when I opened the e-mail, it said that they were, I think, they were in England and they didn’t have enough money to come back to the States so can you send us some money, wire us some money, over, yeah. And if I’m not mistaken, I was probably the first to contact them that they were hacked. I’m like, ‘This isn’t right. Something strange’” (P11)

In other words, a specific negative experience (i.e., receiving odd requests for money from a friend via e-mail), triggered P11 to reach out to this friend to caution that friend about that his email account was likely breached. In other cases, participants relayed cautionary tales to others who were not, themselves, part of the incident. For example, after his girlfriend illicitly accessed his e-mail account, P10 spoke to his friends to let them know that she may have read their conversations:

“It was just like, ‘Hey, [my girlfriend’s] been reading through our mail, like our conversations and stuff,’ [...] She probably read some of our conversations, not like she’s going to get into your accounts.” (P10)

Targeted warnings (7/118 examples)

Another common conversation we observed was one in which a conversation initiator issued someone a warning about potential security or privacy threats after observing that someone engaged in what they believed was risky behavior—what we call **targeted warnings**. For example, one participant described a friend warning her about the danger of not having a passcode:

“I was having a conversation with somebody and they were saying, ‘Don’t you have your passcode on there anymore?’ And I said, ‘No, it’s a pain in the butt.’ And they said, ‘Well, it’d probably be a good idea if y- especially if you like leave it lay around on your desk or something like that. Or even if you’re out in the evening and you have it on your purse, which most people now when they’re out they have this thing right on the table where they are that somebody doesn’t come by and grab it or whatever. That way they can do whatever they want with it.’” (P7)

In this example, P7 conveys a story about how someone told her, after noticing that she no longer used a PIN on her iPhone, that her behavior could result in a number of security breaches.

Spreading the news (8/118 examples)

News articles about security breaches often resulted in conversations we refer to as *spreading the news*—conversations where the initiator attempted to warn friends and loved ones about a security threat outlined in a news article. These conversations sometimes included advice on how to change behavior to protect oneself from the new threat, but were usually just meant to raise awareness that

a threat existed. For example, one participant talked about his contacts on Twitter discussing stories about Facebook privacy concerns without giving advice:

“Oh. Yes. People have said constantly on Twitter about how Facebook, it’s not private anymore. Which is ironic, because neither is Twitter. So I’ve seen that, but no one has showed a article about being secure like with NSA and stuff.” (P4)

As with other warnings, these conversations were often motivated by a desire to protect. For example, one participant described sharing a link to an article, through social media, about a credit card breach in order to warn her loved ones to be careful. Indeed, when asked why she shared one such news article, P2 said:

“To ask my beloved to actually pay attention to these things, to make sure they’re okay. Their bank accounts are okay, if they actually do some shopping that day.” (P2)

Conversations prompted by news articles also sometimes led to sharing best practices or details of privacy and security behaviors.

“We were just generally sitting around and somebody was like, ‘Oh, this is an article about Facebook privacy stuff again. Let’s look at it’ ‘Do you use this,’ or ‘I use that,’ and ‘Oh.’ So really just comparing notes is the best way I can put it. Like we weren’t overly scrutinizing each other’s things. But like ‘I found this to be effective.’” (P10)

In all of these examples, a news article about a potential security breach triggered a conversation between participants and their contacts. In all cases, the purpose of the conversation was primarily to make just others aware of the issue—thus, these conversations were less directed and driven than cautionary tales or targeted warnings.

Teachings

Apart from warnings, the other broad category of conversations we uncovered was *teachings*. Whereas warnings were meant to inform participants about threats, teachings were meant to share security best practices or demonstrate to others how to protect themselves from security and privacy threats. In contrast to warnings, teachings focused on sharing specific information about *behaviors to enact* rather than just information. Another difference was that whereas warning conversations were almost always *reactive* (i.e., triggered by an immediate threat or news about that threat), teachings were both *reactive* (i.e., meant to *solve* an immediate problem) and *proactive* (i.e., meant to *avoid* future threats). Within teachings, we identified three common conversations: *lecturing*, *configuration help*, and *social learning*.

Lectures (8/118 examples)

Lectures occurred when conversation initiators offered security and privacy advice to those for whom they felt accountable—for example, parents and children, or managers and employees. For example, parents often advised young and college-bound children not to over share on Facebook. Older children, however, tended to be the ones lecturing their parents about security best practices. Indeed, when asked if he shared security advice with others, one adult participant said:

“I mean, I’ve spoken to my mom and dad about it. Like, I’ve told them, like, because I’ve told them to also use the same features that I do. Like having screen locks for phones and being more careful about passwords. And not logging into public computers and just leaving them without signing out.” (P8)

P8’s advice to his parents is a good example of *proactive* advice—advice about security given even in the absence of an immediate threat. This example is also illustrative of another theme that often

arose with teachings: a sense of *accountability*. P8 felt accountable for the security of his parents, which in turn prompted these conversations.

Another type of lecture was managers lecturing employees about best practices to protect company data. For example, when asked why she updated her e-mail password, one participant said:

“Actually, this [advice] was given to me by my manager, with whom I used to work. So he’s the one who told me about this. He was like you should change your password because it contains confidential information.” (P13)

Similarly, another participant described his boss asking him to encrypt confidential files and transmit them physically on a USB flash drive rather than through email (P18).

Configuration help (8/118 examples)

Often, teaching conversations were triggered by a conversation initiator soliciting advice on how to configure security and privacy settings for a new device or account—what we called *configuration help* conversations. For example, one participant described helping his mother set up her new laptop with the appropriate security settings to keep her information safe (P19). Another participant described encouraging his mother to enable 9-dot authentication on her new Android phone to make sure no one else could access it. When asked why, P15 responded:

“I mean, just the same reason that people shouldn’t just look into her phone. Because, like, if it does not have a button, anyone can just, like, unlock and look at her messages and stuff.” (P15)

While not directly the trigger for this conversation, this excerpt is again illustrative of the sense of accountability that people feel for their loved ones’ security and privacy.

Broadly, configuration help conversations were about setting up Facebook privacy settings (P1, P3, P4, P8, P19). For example, P19 describes how her mom asked for her help with enacting specific privacy related behaviors on Facebook.

“my mom...doesn’t really know how to do Facebook that much so she’ll ask me questions about it, in general, like how to post or, I guess, how to remove herself from something or certain things like that. So, I guess, I have given her advice in a way, just given her a few basic steps of set this as this just so you don’t have-- you’re not completely open and public.” (P19)

Social learning (5/118 examples)

Finally, while they were not very frequent among our participants, one of the more effective conversations at enacting behavior changes we uncovered were *social learning* conversations. In these conversations, conversation initiators observed novel security or privacy behaviors or tools and asked those who enacted those behaviors or used those tools questions about those behaviors or tools. For example, some participants asked others about novel ways to construct passwords (P9, P10, P18) or a new type of authentication (P8). For example, P18 asked a friend about sharing his Amazon account password, prompting the friend to share his password composition method:

“When I was working this summer, one of my co-workers told me about the whole algorithm thing. One, it just helps you I guess have different passwords. It helps you recall them easier based on I guess the type of profile. I guess you can cater, you can change your algorithm, depending on I guess what you want to be in it. But ever since I started using it.” (P18)

Typically, our more security-savvy participants reported that they did not often share their own security behaviors with their less security-savvy friends and loved ones because they feared the topic was too boring. However, social learning conversations presented opportunities for experts or early adopters to share their solutions for solving common security problems.

Discussion

In analyzing the 118 conversations about security and privacy reported by our participants, I uncovered six common conversation catalysts (see Table 5. Conversation catalysts derived from our iterative open coding process.) and seven common conversation goals (see Table 6). From these catalysts and goals, I constructed an initial typology for conversations about security and privacy in which I identified six of the more common and interesting catalyst-goal contexts (see Table 8). In turn, this typology affords an answer to my research question: *under what circumstances do people generally talk about privacy and security?*

Broadly, the answer to this question appears to be: to *warn* or to *teach* others. Most commonly, my participants reported conversations about privacy and security to be educational experiences—either in sharing or receiving information about a novel security threat, or in sharing or receiving advice about how to solve a specific security problem or security best practices. *Observability*, again, *appeared to be a key driver of conversations*—be it experts witnessing *insecure* behavior or non-experts witnessing *novel* behavior. These findings reaffirm the notion that social processes contribute to the modulating of security sensitivity, as these conversations often raised awareness, motivation and/or knowledge about security threats and behaviors. Specifically, *warning conversations* typically raised awareness and motivation, while *teaching conversations* typically raised awareness and knowledge.

Opportunities

Create more opportunities for social learning. One type of conversation that raised all parts of the security sensitivity stack, *social learning*, was not very prevalent despite its efficacy. Indeed, social learning conversations may represent the ideal context under which social influence *can* affect security sensitivity—novices interested in learning about security voluntarily ask for information from experts, thereby raising their awareness, motivation *and* knowledge. In turn, experts are willing to share their information and don't feel that their efforts are wasted, as was implied by two expert participants when asked why they don't share information about threats more often (P4, P9). One way to increase opportunities for social learning may be to increase the *observability* of security tools and behaviors. Indeed, the few social learning conversations that *did* occur were triggered when people saw novel security behaviors or tools in practice (e.g., one participant reported being asked, by a stranger, about Android 9-dot authentication at a café because it looked interesting).

Facilitate experts' sharing of security advice with others. Unfortunately, many of my participants alluded to an illusory correlation [20] between security feature usage and paranoia, referring to their expert friends as “hyper-secure” (P5) and their actions as going “above and beyond” (P18) or “nutty” (P1). Perhaps as a result of this negative perception towards those with high security sensitivity, many of the security savvy participants I interviewed mentioned that they avoided sharing *proactive* information with their friends because the topic seemed socially inappropriate or unwelcome—as too “preachy”, for example. There is, thus, a substantial missed opportunity for experts to share knowledge with novices that only appears to be overcome when novices observe and query about interesting, novel behavior by the expert. Apart from increasing the observability of security tools and behaviors to create more opportunities for social learning conversations, another way to facilitate experts' sharing of security advice might be to create tools for experts to share their knowledge *anonymously* (e.g., so they can share this information without incurring any perceived losses to social capital) or *indirectly* (e.g., by generally sharing their behaviors for others to view at their own leisure, but not as a directed message at any specific time).

Create security tools that allow people to act on their sense of accountability. Many conversations about security and privacy were triggered by a sense of accountability participants felt for the security of their loved ones. Yet, few security and privacy tools exist that allow participants to act on this sense of accountability—for example, tools that allow people to audit their loved ones security and privacy configurations. From the literature on usable privacy and security, we know that it is difficult to get people motivated enough to enact security and privacy behavior changes for their own sake [56]. My results indicate, however, that it may be easier to get people motivated enough to help their loved ones enact positive security and privacy behavior changes.

Chapter 5: How Social Influences Affect Security Tool Diffusion

The contents of this chapter are drawn from a previously published paper: [The Role of Social Influence on Security Feature Adoption](#). Sauvik Das, Adam Kramer, Laura Dabbish and Jason Hong. In Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'15) [30].

Summary

The interview study results offered initial evidence that social influence affects security behaviors. However, it had two limitations: (i) it was limited in its scale with only 19 participants, and (ii) it relied upon participants' memory of their behaviors. In this chapter, to get a larger-scale understanding of how social influence affects security behavior, I report on a large scale investigation of how three optional security tools diffused through the social networks of 1.5 million Facebook users. My results confirm the findings that social influence can affect security tool adoption both positively and negatively. Specifically, the directionality and magnitude of the effect of social influence on any potential adopter's likelihood of adopting a security tool is modulated at least by: (i) the number of the potential adopter's friends who currently use the tool and (ii) the design of the security tool itself. Indeed, when one is exposed to relatively few friends who use a security tool, negative social proof appears to create a disaffiliation effect that inhibits the tool from further adoption. Conversely, when one is exposed to relatively many friends who use a security tool, positive social proof appears to increasingly encourage further adoptions. Furthermore, security tools that are *observable*, *socially inclusive*, and allow for *stewardship* are far more amenable to social spread. Taken together, these results offer an explanation for why many security tools are rarely used despite improvements to usability. Specifically, as most security tools are designed to be unobservable and socially exclusive, the early adopters of these tools tend to be those who care especially about security. Lay users may perceive these early adopters as being "paranoid" or caring too much about security. Accordingly, it is possible that the early adopters of security tools create a disaffiliation effect which inhibits further adoption outside of a core group of security enthusiasts.

Motivation

In the previous chapters, I reported on a rich, albeit small-scale interview study in which we showed that social influence plays a pivotal role in security related behavior change and security tool adoption. However, our results suggested that social proof can have contradictory effects.

Sometimes social proof appears to promote adoption of security technology and behaviors—especially when the tools and/or behaviors being diffused are highly novel or observable [11]. Indeed, as we know from a rich body of prior work in social psychology research, *social proof* can be an effective motivator for behavior change. This holds true even for behaviors, like security behaviors, that are not immediately or even directly gratifying—for example, conserving energy, reusing hotel towels, or preserving natural parks.

At other times, social proof appears to stifle the adoption of security technology because security tools tend to be preventative, intrusive, and associated with paranoia [28,45]. Indeed, if only 'experts' or people who are perceived as paranoid initially use a security tool, lay people might develop an illusory correlation [20] between using a security tool and paranoia that disenfranchises the use of a security tool.

Taken together, it appears that social influence can be both a helpful and harmful force in security-tool adoption, but we do not yet fully understand the parameters under which it is helpful or harmful. In addition, we do not know how social influence plays out unadulterated "in the wild"—the examples reported by our interview participants are subject to recall bias where only especially memorable instances were reported. Thus, to have a clearer understanding of how social forces affect security tool adoption, as well as to understand when positive social proof drives adoption and when negative social proof stifles adoption, I next analyzed how the adoption of three Facebook security tools—Login Notifications, Login Approvals and Trusted Contacts—diffused through the social networks of 1.5 million people [2]. The specific research questions we were trying to answer with this work were:

Q1. In practice, does social influence have a detectable effect on one's likelihood to adopt a security tool?

Q2. If so, what factors affect the directionality and magnitude of the effect of social influence on security tool adoption?

Hypotheses

With big data observational research, it is easy to find significant effect sizes and retrofit hypotheses after the results are known. To avoid falling into this trap and better frame our methodology, we first surveyed the prior literature in social psychology, social network analysis and usable security in order to develop believable hypotheses.

Social Diffusion and Friend Diversity

Earlier work suggests that exposure to novel information on social networking sites increases information diffusion through social channels [7], but that these diffusion chains are most effective when the seed information is shared by many different sources [7,101], especially when the information is intended to enact behavior change [18,19]. Ugander and colleagues [106] extended this result, finding that people who were invited to join Facebook through e-mail recommendations from their friends were more likely to join if the recommenders were from distinct social contexts—i.e., receiving an invitation from a school friend and a family member was more convincing than receiving invitations from two different family members. Romero and colleagues [89] found that the “persistence” of the information being spread—or, the marginal likelihood that content will be re-shared after one more exposure—is also important in determining whether content will be diffused. Specifically, controversial topics—like information about security, say—require repeated exposure from many sources before they are diffused. These considerations led us to hypothesize:

H1: People with exposure to tool-adopting friends from many distinct social contexts will be more likely to use that tool than others with exposure to the same number of tool-adopting friends from fewer distinct social contexts.

Social Diffusion and Observability

It is well established that not all behavior diffuses equally [18,19], and the adoption of technology is no different. Thus, efforts have been made to model the factors that influence the adoption of technology. Rogers [86], in his seminal work on the diffusion of innovations, argued that new technology gets widely adopted through a process by which it is communicated through members of a social network. Rogers argues that primarily subjective perceptions get communicated through social channels, and that these perceptions are key to the success of an innovation. He further outlines that preventative innovations—or innovations, like security tools, that prevent undesirable outcomes from happening—typically have low adoption rates, in part because of their low *observability*, or the invisibility of their use and benefits. Finally, in my own prior work reported in Chapter 3: A Typology of How Social Influence Affects Security Behaviors and Chapter 4: Understanding How Security Information Is Communicated, I found that the *observability* of security tools and behaviors was a key factor in driving the adoption of security tools. In fact, I found that of all social catalysts for behavior change, observing others use security tools was the most prevalent.

H2: More observable security tools will more effectively diffuse through social channels than less observable security tools.

Social Diffusion on Security Adoptions

Prior work in psychology and the application of social influence implies that if many of one's friends and acquaintances use a security tool, one should be more likely to use that security tool herself. This is the basic premise of social proof—that we look to each others for cues on how to behave when we are uncertain [21]. Yet, we see some counter examples of this premise in the usable security literature. Indeed, Gaw and colleagues [45] found that many non-experts perceived others who used e-mail encryption as “paranoid”, a perception that inhibited their own use of e-mail encryption. In

our interview study, we found that non-expert participants were similarly aversive towards using security tools, and spoke of their security-expert friends as being “nutty” or going “above and beyond”.

Thus, it appears that social proof does not always have the expected effect on security tool adoption. We believe, in fact, that because security tool usage is often invisible, rarely communicated, and generally undesired [55,92], social proof can act *against* the adoption of a security tool at early stages of adoption.

Indeed, prior work in usable privacy and security suggests that many security tools remain unused because stringent security measures are often antagonistic towards the specific goal of the end user at any given moment [92]. For example, while a user might want to check her e-mail, a complex password that usually requires three attempts to get right *prevents* her from checking her e-mail. Thus, people often reject security tools when they expect or experience them to be weighty [4]. Consequently, typically only people who are especially dedicated to protecting their information use interruptive security tools, and we know from prior work that non-experts may perceive these early adopters as “paranoid” [45]. More formally, because early adopters of security tools are likely to be perceived by others as behaviorally *different* (e.g., either paranoid, or in possession of expert knowledge), non-experts may perceive an illusory correlation [20], or an exaggerated relationship, between security tool usage and this behavioral difference.

In turn, as non-experts consider themselves different from those who use security tools, they may reject the use of security tools. Moreover, this illusory correlation should only *strengthen* as more of these security-enthusiast early adopters use the tool because of the “availability heuristic”—a mental shortcut that biases people’s judgments towards what is more frequently recalled [105].

The upshot is that the subjective perceptions of a security tool that propagates through social channels may be tainted into working *against* its adoption, at least until enough of a potential adopter’s behaviorally similar friends start using the tool so that its use becomes normative.

In other words, there may be a non-linear relationship between one’s exposure to tool-adopting friends and one’s likelihood to adopt a security tool. Specifically, if a potential adopter is only exposed to *few*, early-adopter friends who use a security tool, it is possible that he might find social proof that a security tool should *not* be used (because of an illusory correlation), and the strength of this *negative* social proof should increase with the number of these tool-adopting friends (because of the availability heuristic). On the other hand, once a potential adopter is exposed to *many* tool-adopting friends, especially those that are similar to himself, he might find social proof that a security tool *should* be used (because of the positive effects of homophilous networks on technology adoption [8]), and the strength of this *positive* social proof should increase with the number of his tool-adopting friends.

H3: When a potential adopter is exposed to many tool-adopting friends, he will be more likely to adopt a security tool than those with fewer tool-adopting friends.

H4: When a potential adopter is exposed to few tool-adopting friends, he will be less likely to adopt a security tool than those with even fewer tool-adopting friends.

Methodology

In the summer of 2013, I was intern on the Facebook Data Science team³, which afforded me access to the anonymized and aggregated security tool adoption patterns of Facebook users. To test our hypotheses, we monitored security tool adoptions for the following three Facebook security tools: (1) **Login Approvals**—A tool that requires adopters to enter a separate code, usually generated on

³ <https://www.facebook.com/data>

or sent to the adopter’s smartphone, in addition to their password when they attempt to authenticate; (2) **Login Notifications**—A tool that notifies adopters, via e-mail or SMS, when their account is accessed from previously unseen browsers and devices; and, (3) **Trusted Contacts**—A tool that allows an adopter to specify three to five friends who can verify her identity if she forgot her password and cannot access her e-mail. We investigated multiple security tools to avoid drawing conclusions specific to any one tool, as well as to empirically evaluate the hypothesis that the design of a security tool (e.g., its observability) may play a role in its diffusion [21]. Furthermore, we chose these three tools because of their diversity and colocation within the “security settings” page on Facebook.

For 12 days in late 2013, we collected data from a random subset of people who use Facebook and newly adopted one of the aforementioned security tools: Login Approvals, Login Notifications, or Trusted Contacts. In total, we collected data from n=250,000 people per tool (750,000 adopters overall)—the positive examples of tool adopters in our dataset. Then, for each day and tool, we also obtained a random sample of an equal number of people who had not adopted that tool up to that day—negative examples of tool adopters. In total, we had n=1,500,000 people across all twelve days, three tools (Login Approvals, Login Notifications, Trusted Contacts), and two tool usage states (i.e., uses or doesn’t use).

For all people in our sample, we also collected a set of variables that we believed could have affected one’s decision to adopt a security tool. These variables fell under four categories: demographic variables that described individual characteristics such as age and gender; behavioral variables that described activity on Facebook, such as posts shared and deleted; network variables that described one’s social network, such as friends’ average age and gender diversity; and, social proof variables that described how many and which of a person’s friends had adopted any of the aforementioned security tools up to the day during which the data was collected. In Table 9, we provide a full list of variables included in our analysis. All data was de-identified prior to our analysis.

Demographic Variables	
<i>Age</i>	Age of the individual.
<i>Gender</i>	Self-reported gender: male or female.
<i>Friend count</i>	Count of the individuals number of friends with Facebook accounts.
<i>Account length</i>	Days that have passed since the individual activated his account.
<i>Days active in last 30</i>	Days the individual was active on Facebook in the past 30 days.
Social Network Variables	
<i>Mean friend age</i>	Average age of the individual’s Facebook friends.
<i>Friend age entropy</i>	Shannon entropy of the individual’s Facebook friends’ ages.
<i>Percent male friends</i>	Percentage of the individual friends that are male.
<i>Mean friends’ account length</i>	Average number of days an individual’s Facebook friends have used Facebook.
<i>Friend country entropy</i>	Shannon entropy of countries from which the user has friends.
<i>Mean number of friends among friends</i>	Average number of Facebook friends among an individual’s Facebook friends.
Behavioral Variables (all aggregated across the week prior to data collection)	
<i>Posts Created</i>	Number of posts created.
<i>Posts Deleted</i>	Number of posts deleted.
<i>Comments Created</i>	Number of comments created.
<i>Comments Deleted</i>	Number of comments deleted.
<i>Likes</i>	Number of likes given.
<i>Friends Added</i>	Number of friends added.
<i>Friends Removed</i>	Number of friends removed.
<i>Photos Added</i>	Number of photos added.
<i>Videos Added</i>	Number of videos added.
Social Proof Variables	
<i>Percent of friends who use Login Approvals</i>	Percent of friends who use the Login Approvals security tool.
<i>Percent of friends who use Login Notifications</i>	Percent of friends who use the Login Notifications security tool.
<i>Percent of friends who use Trusted Contacts</i>	Percent of friends who use the Trusted Contacts security tool.
<i>Number of diverse social contexts</i>	Number of social contexts from which friends who use security tools originate.

Table 9. Collected tool descriptions. These variables were all collected per individual.

We selected people who newly adopted security tools because security tool adoptions were not time-stamped in our data, so it would be otherwise impossible to know who, between two people, adopted a security tool first. For someone who newly adopted a security tool on a given day, however, we knew that all friends of their friends who used that tool adopted it before that day.

Notably, we could not measure how security tool adoptions diffused—i.e., we did not alter the observability of security tool usage and initiation. Rather, we simply control for other factors that also affect security tool adoption, such that we can compare the tool adoption rate of two sub-populations that differ primarily in their exposure to friends who have adopted a security tool. We do not believe this limitation to be stifling—understanding the channels through which social diffusion occurs is separate from our goal of understanding its ultimate effect on security tool adoption.

Finally, all data collection complied with Facebook’s terms of use and data use policy and was performed in aggregate so that we were not privy to any individual’s information. Furthermore, as our data was observational, we believe our analysis constituted minimal risk to those in our sample.

Results

How exposure to friends from distinct social circles relates to adoption

Analysis method

First, we wanted to test the hypothesis that people with security-tool adopting friends from many distinct social circles should be more likely to adopt a security tool than those with the same number of tool-adopting friends from fewer distinct social circles. To do so, we estimated a logistic regression model for each security tool. These regressions modeled the strength of the relationship between a person’s likelihood to adopt a security tool and the number of distinct social contexts from which his tool-adopting friends originated. Note that we define a “distinct social context” as a distinct connected component in one’s friend graph, following similar definitions used in prior work [29].

As linear regression analysis assumes independence in the response variable (in our case, whether or not someone in our sample adopted a security tool), we only included a balanced subset of our full sample into the regressions after eliminating people in our sample who happened to be Facebook friends with one another. This reduced sample consisted of $n=65,000$ positive and negative examples of tool adopters for each of our three tools, resulting in $n=130,000$ people for each regression, all of whom were not friends with one another.

In running these regressions, we controlled for the demographic, social network and behavioral variables described in Table 9. In addition, we also controlled for the *number* of one’s tool-adopting friends, so that the coefficient for the *number of distinct social contexts* variable can be interpreted after controlling for a potential adopter’s number of tool using friends.

Analysis Results

The coefficient for the *number of distinct social contexts* variable for each logistic regression is shown in Table 10. These coefficients represent a change in “log-odds”, or $\ln \frac{P}{1-P}$, where P represents the probability that an individual adopted the security tool. A positive coefficient implies that the log-odds ratio increases, or that an increase in the variable increases the likelihood that a person adopts the tool, P . A negative coefficient implies the opposite. Furthermore, each variable was centered and scaled, such that its coefficient represents the expected change in log-odds that a person uses a tool given a one standard deviation increase in the predictor variable, holding all other numerical variables at their means and categorical variables at their baselines. Additionally, larger absolute coefficient values imply a stronger relationship between the IVs and DVs.

Thus, from Figure 4 and Table 10, we can see that the *number of diverse social contexts* variable positively correlated with the adoption of *every* security tool ($b_{LA}=+0.15$, $p<2e-16$; $b_{LN}=+0.03$, $p<2e-16$; $b_{TC}=+0.88$, $p<2e-16$). This finding offers support for **H1**—people with friends from more diverse social contexts (e.g., high school friends, college friends, family) who use a security tool should be more likely to adopt that tool themselves than those with tool-adopting friends from fewer distinct

social contexts. In other words, it is not just the *number* of one’s friends who use a security tool that

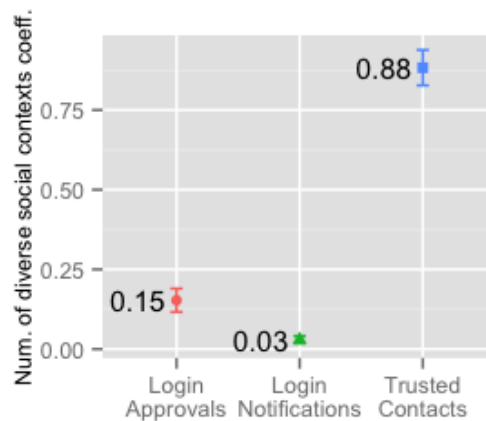


Figure 4. Coefficients for the three logistic regressions relating the number of diverse social contexts variable to use of each security tool, with 95% confidence intervals. All coefficients significant, $p < 2e-16$.

Variable Name	Login Approvals	Login Notifications	Trusted Contacts
Intercept	0.28 ***	0.12 ***	0.31 ***
Age	-0.06 ***	0.08 ***	-0.01
Gender: male (relative to female)	0.05 ***	-0.06 ***	-0.10 ***
Days with active account	-0.04 **	-0.26 ***	0.02
Friend count	-0.07 ***	-0.11 ***	-0.03 ***
Days active in past 30	0.62 ***	0.58 ***	0.50 ***
Mean friend age	-0.47 ***	-0.21 ***	-0.49 ***
Friend age entropy	-0.16 ***	-0.36 ***	-0.07 ***
Percent male friends	0.36 ***	0.34 ***	0.43 ***
Mean friends’ days with active account	-0.84 ***	-1.00 ***	-1.04 ***
Friend country entropy	0.32 ***	0.21 ***	0.29 ***
Mean number of friends of friends	-0.08 ***	-0.02 **	-0.14 ***
Posts created	-0.20 ***	0.19 ***	-0.17 ***
Posts deleted	0.27 ***	0.20 ***	0.15 ***
Comments created	0.10 ***	0.06 ***	0.18 ***
Comments deleted	0.18 ***	0.17 ***	0.16 ***
Likes given	-0.07 ***	-0.09 ***	-0.01
Friends added	1.81 ***	2.37 ***	1.36 ***
Friends removed	0.57 ***	0.49 ***	0.50 ***
Photos added	0.10 ***	0.14 ***	0.26 ***
Videos added	-0.01 ***	-0.02 **	-0.02 **
Percent of friends who use feature	0.13 ***	-0.12 ***	0.29 ***
Number of diverse social contexts	0.15 ***	0.03 ***	0.88 ***

Table 10. Coefficients for the three logistic regressions relating social proof variables (bolded, at the bottom), to use of login approvals (left), login notifications (middle) and trusted contacts (right). All coefficients are normalized.

*** $p < 2e-16$, ** $p < 0.001$

matters; these friends should be independent of one another for the effect to be strongest.

In addition, the discrepancy of effect size across tools offers some support for **H2**—that more observable security tools will be more effectively diffused through social channels. Indeed, the

absolute effect size of the *number of diverse social contexts* variable is largest, by far, for Trusted Contacts (the most observable tool, $b_{TC}=+0.88$), then for Login Approvals (the next most observable, $b_{LA}=+0.15$) and finally lowest for Login Notifications (the least observable tool, $b_{LN}=+0.03$).

Indeed, Login Notifications are private messages that are not very observable, and are thus difficult to passively diffuse via social channels. Thus, while having many different friends use Login Notifications may make for a more convincing case for a potential adopter to use the tool, the case is unlikely to be made. Login Approvals are more observable than Login Notifications in that friends who are collocated with an adopter will see the additional authentication step it requires, which in turn may passively provide these friends with social proof to use Login Approvals [28]. This modest increase in observability appears to correlate with a modest increase in the effect size of the *number of diverse social contexts* variable. Finally, the Trusted Contacts tool sends out a notification to each of one's friends who was specified as a Trusted Contact, thus substantially increasing its visibility in a direct way and, in turn, correlating with a substantial increase in effect size. It is also possible that the social nature of the tool—in enlisting friends to help recover one's account—lends itself to amplified social diffusion.

In summary, our regression analysis provides us with support for **H1** and limited support for **H2**, but we have yet to test **H3** and **H4**—that the tool-adoption rate of one's current set of friends will moderate whether the effect of social proof will be positive or negative on one's own likelihood to adopt that security tool. Unfortunately, linear regression analysis is limited in that it does not consider this form of non-linearity in the relationship between predictor and response. Furthermore, regression analysis confounds homophily-based diffusion with social-influence based diffusion [6,97]. In other words, because similar people cluster together as friends, we cannot tell if co-adoption of a tool is due to one friend influencing another or because both friends share an interest.

How social influence affects the adoption of security tools

Analysis Methodology

Thus, to test **H3** and **H4**, we ran an adapted version of *matched propensity sampling* analysis [6]. Matched propensity sampling is a form of causal inference that helps us differentiate tool adoption due to homophily from tool adoption due to social influence. It distinguishes between homophily and social influence by comparing the tool adoption rates of two sets of people who are *equally likely* to have a fixed proportion of friends who have adopted a security tool, where one set *actually does* have the fixed proportion of friends who have adopted this tool and the other set does not. People in the former set are “exposed” to their tool-adopting friends at this fixed rate, while those in the latter set are “unexposed.”

Exposed and unexposed individuals are matched, in pairs, based on a “propensity score” computed from a set of covariates Z that are theorized to represent homophily-based diffusion [90]. We used a logistic regression to calculate the propensity score as suggested by prior work [6], and the covariates included in the model were the demographic, behavioral, and social network variables listed in Table 9. As we are not concerned about estimating exact coefficients and their variances with the logistic regressions in this analysis, we are able to break the independence assumption and include the full set of 1.5 million users in our sample.

Unfortunately, as we could not capture the security expertise of those in our sample, there remains some form of “latent homophily” for which we do not control. However, the demographic, behavioral, and social network variables for which we control likely predict security expertise, so we believe this limitation to be minimal.

By matching exposed and unexposed individuals who have the same likelihood of being exposed, we can take the difference in tool adoption rates between the exposed and the unexposed as evidence of the effect of social influence. Indeed, after the propensity matching process, the only theoretical difference between these two sets of people are that the exposed set has a certain proportion of friends who use a security tool and those in the unexposed set do not. If social influence has no effect, we should see the *same* rate of adoption for the exposed and unexposed, whereas if social influence

has a positive or negative effect, we should see that exposed individuals adopt the tool at a *higher* or *lower* rate, respectively.

We specified five empirical exposure conditions for each security tool—Login Approvals, Login Notifications, and Trusted Contacts—with each exposure condition representing whether or not the user was at least in the 1st percentile, the 21st percentile, the 41st percentile, the 61st percentile, or the 81st percentile in the *percent of friends who use tool* variable, or the total percentage of their friends who used a security tool at the day of data collection. Notably, a potential adopter could count as “exposed” at some levels but not others.

	Percentile	Approvals	Notifications	Trusted Contacts
<i>E1</i>	1 st	0.2%	2.0%	0.1%
<i>E2</i>	21 st	0.8%	7.3%	0.4%
<i>E3</i>	41 st	1.3%	10.0%	0.7%
<i>E4</i>	61 st	1.8%	12.3%	1.1%
<i>E5</i>	81 st	2.7%	15.1%	2.0%

Table 11. Exposed condition prerequisites for each security tool. For example, if a user is “exposed” at *E3* for login approvals, at least 1.3% of her friends must have adopted login approvals at the time of data collection.

Figure 5 depicts the values of the *percent of friends who use tool* variable that qualified for “exposure” under *E1* through *E5*, with actualized values for these conditions shown in Table 11. Concretely, an individual is exposed in *E1* for Login Approvals if at least 0.2% of her friends adopted the tool, because that puts her at least at the 1st percentile of people whose friends have adopted the tool.

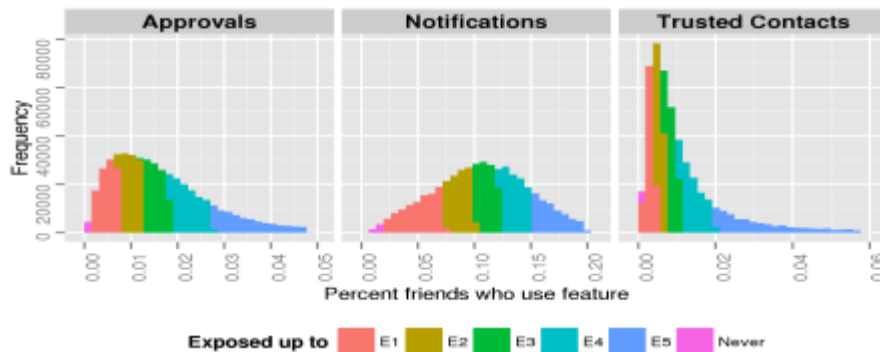


Figure 5. Histogram of percent of friends who use login approvals (left), login notifications (middle) and trusted contacts (right). Colors represent up to what exposed conditions users with *x%* of tool-adopting friends would be considered “exposed” in the analysis

Likewise, she is exposed in *E5* for Login Approvals if at least 2.7% of her friends adopted the tool.

We chose five exposure conditions uniformly spaced across the distribution of the *percent of friends who use tool* variable to get a detailed map of the relationship between exposure to friends who have adopted a security tool and one’s own likelihood to adopt that tool at different levels of exposure. This map should help us evaluate both **H3** and **H4**—specifically, **H3** predicts a higher adoption rate for the *exposed* relative to the *unexposed* at high exposure conditions because of positive social proof, whereas **H4** predicts a higher adoption rate for the *unexposed* at low exposure conditions because of negative social proof—specifically, an illusory correlation between the attributes of early adopters (e.g., “paranoid”, “nutty”, “expert”) and the security tool itself.

Analysis Results

Figure 6 shows the rate of feature adoption for exposed and unexposed individuals for all three features across all five exposures. In interpreting the results of the matched propensity analysis in Figure 6, we note the following: (i) If social influence has any effect on the adoption of a security

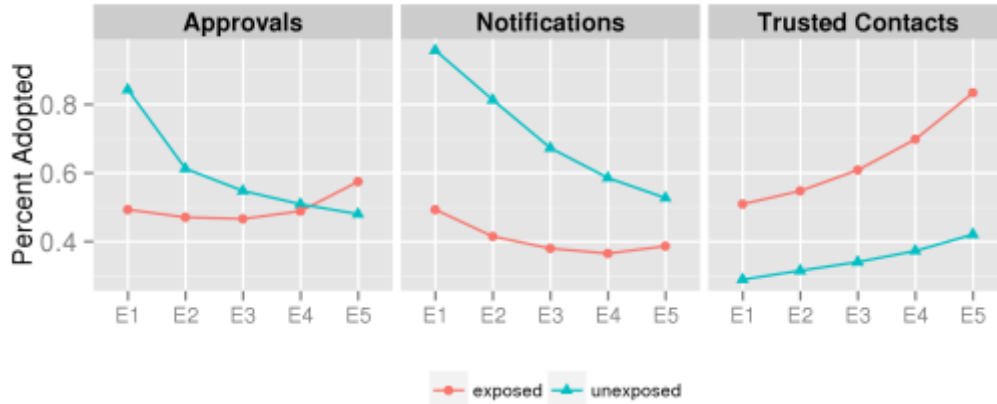


Figure 6. Feature adoption rates, plotted for each security feature for each exposure condition, for both exposed and unexposed individuals. Exposed feature adoption rates are plotted as red circles, and unexposed feature adoption rates are plotted as blue triangles.

feature at a particular level of exposure, we should see a significant difference in the adoption rates of exposed and unexposed individuals; (ii) If social influence has a positive effect on the adoption of a security feature at a particular level of exposure, then we should see that exposed individuals have a significantly higher adoption rate than the unexposed; and, (iii) If social influence has a negative effect on the adoption of a security feature at a particular level of exposure, we should see that exposed individuals have a significantly lower adoption rate than the unexposed.

	<i>Approvals</i>		<i>Notifications</i>		<i>Trusted Contacts</i>	
	N	χ^2, df=1	N	χ^2, df=1	N	χ^2, df=1
<i>E1</i>	5852	1553	25061	13743	4995	491
<i>E2</i>	122765	4994	518907	172603	105156	11742
<i>E3</i>	240061	3104	1014159	174619	205541	29775
<i>E4</i>	228905	140	963824	93771	196397	42022
<i>E5</i>	111092	1976	468147	18828	95393	34665

Table 12. Chi square significance tests for the difference in adoption rate between exposed and unexposed individuals across all exposure conditions and all security features. All differences significant, $p < 2e-16$.

First, as we show in Table 12, below, all of the differences in adoption rate between the exposed and unexposed were significant, suggesting that irrespective of the security feature and level of exposure to friends who use that feature, social influence appears to have a significant effect on one's likelihood to adopt a security feature. This finding strongly supports our smaller-scale qualitative results, explained in Chapters Chapter 3: A Typology of How Social Influence Affects Security Behaviors and Chapter 4: Understanding How Security Information Is Communicated, that surfaced social influence as a key factor in the adoption of security features [28].

For Login Notifications, we see that people who are exposed to a certain proportion of feature-using friends appear to be less likely to adopt those features than people who are unexposed for all levels of exposure we tested. Thus, in our sample, even people with a higher-than-average proportion of feature-adopting friends (i.e., those exposed at E4-E5 who are at least at the 61st percentile) were themselves less likely to use Login Notifications than people who had fewer friends who used those features. It appears, therefore, that exposure to friends who use Login Notifications stifles the adoption of Login Notifications, a finding that supports **H4**—that social influence will have a negative effect on feature adoption at low exposure levels—but conflicts with **H3**—that social influence will have a positive effect on feature adoption at high exposure levels.

We see just the opposite trend for Trusted Contacts, however: even at E1, the lowest level of exposure, exposed individuals are significantly more likely to adopt Trusted Contacts than the unexposed. In other words, it seems that any exposure to friends who use Trusted Contacts substantially increases one’s own likelihood to adopt that feature, a finding that supports **H3** but contradicts **H4**.

Finally, for Login Approvals, we see exactly the nuanced, thresholded relationship we predicted. At lower levels of exposure, unexposed individuals are more likely than the exposed to adopt the feature, but at the highest level of exposure, exposed individuals are more likely to adopt the feature—a finding that supports *both* **H3** and **H4**.

Thus, we have three security features for which adoption is significantly affected by social influence, but for which the effect of social influence appears to manifest differently. For Login Notifications, it appears that social influence is a categorically negative force on its adoption, for Trusted Contacts it is a categorically positive force, and for Login Approvals, the direction of its effect is based on a threshold level of exposure a potential adopter has to friends who already use that feature. What could explain the differences in the effect of social influence across these features?

Theoretical vs. Empirical Exposure Threshold

The matched propensity sampling analysis only reflects the effect of social influence on the adoption of a feature at its rate of adoption at the time of data collection. Indeed, our exposure conditions were based on an empirical division of the percent of friends who use feature variable; therefore, it is possible that there is a theoretical exposure greater than E5 where social influence could have a positive effect on the adoption of Login Notifications. Indeed, for Login Notifications, exposure at E5—at which about 15% of one’s friends use Login Notifications—may not yet be at the threshold where **H3** predicts social influence should have a positive effect on its adoption.

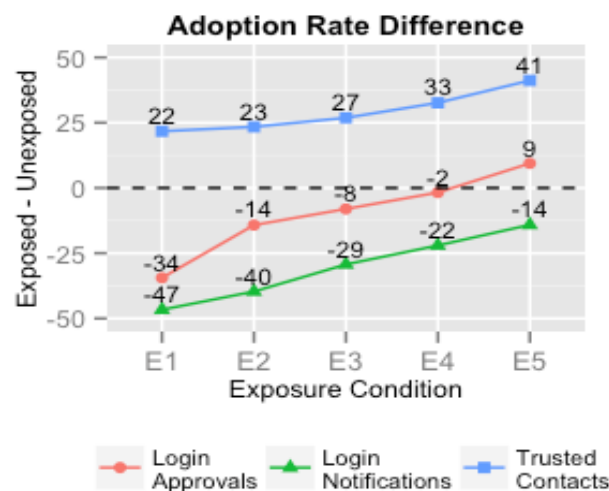


Figure 7. Differences in adoption rate between the exposed and unexposed for all three features across all exposure conditions. Values above the dashed horizontal line signify that those who were exposed had a higher adoption rate than the unexposed. All differences were significant at $p < 2e-16$

To test this possibility, we must observe how the adoption rate difference between the exposed and unexposed varies across exposure conditions. We plot these differences in Figure 7, by subtracting the unexposed adoption rate from the exposed adoption rate. From this plot, we can understand the marginal effect of social influence on adoption at higher exposure conditions. In interpreting Figure 7, we note the following: (i) If unexposed individuals are more likely than the exposed to adopt a feature at a certain level of exposure, then the value of the difference will be negative, whereas it will be positive if exposed individuals are more likely to adopt the feature than the unexposed; and, (ii) If the value of the difference increases (moves up) at higher exposure conditions, then the marginal

effect of having more friends who use a security feature on that feature's adoption is positive, *whereas if the value of the difference decreases (moves down), then the marginal effect is negative.*

From Figure 7, we see that the value of the difference between exposed and unexposed adoption rates increases (moves up) constantly, for all three features, from E1 to E5. For Login Notifications and Login Approvals, the initial adoption rate advantage of unexposed individuals gradually diminishes at higher levels of exposure. In fact, the advantage is ultimately in favor of exposed individuals for Login Approvals at E5, when the difference shifts from negative to positive. For Trusted Contacts, the advantage starts with exposed individuals and simply gets larger at higher levels of exposure. Thus, at higher levels of exposure, the likelihood for exposed individuals to adopt any of the security features grows at a rate faster than the unexposed. It seems likely, therefore, that there is a theoretical exposure higher than E5 where exposed individuals are more likely to adopt Login Notifications than the unexposed—as would be predicted by **H3**. Unfortunately, we did not have a large enough number of people at a high enough exposure to empirically confirm this prediction from the data in our random sample.

It is tempting to also apply this logic to entertain a theoretical exposure lower than E1 at which the effect of social influence is negative for Trusted Contacts. However, as the exposure threshold for E1 for Trusted Contacts is just 0.1%, the theoretical and empirical exposure lower bounds are essentially the same—i.e., having at least one friend who uses the feature. Thus, while it seems like **H3** may be true even for Login Notifications, it seems likely that **H4** may not be true for some features—social influence does not have to be a negative force at low exposure conditions.

Individual Feature Attributes

Another consideration in interpreting the differences in the effect of social influence across security features is the individual attributes of each feature. Specifically, as **H2** suggests, more observable security features should be more positively affected by social influence.

The threshold beyond which the effect of social influence toggles from negative to positive appears to be inversely proportional to the observability of the feature, lending further support for **H2**. Indeed, the threshold is “lowest” for Trusted Contacts in that the threshold seems to be at its theoretical lowest possible value of having just one friend who uses the feature. The threshold is next lowest for Login Approvals at E5—or when approximately 2.7% of ones friends use the security feature. Finally, the threshold is highest for Login Notifications at a level of exposure higher than E5, if such a threshold exists at all.

It makes intuitive sense that the threshold of friends required for negative social proof to be overcome by positive social proof should be lower for more observable features. If our reasoning for **H4** is correct, negative social proof is the result of stereotypes and generalizations that may be overcome if potential adopters can see, concretely, that security feature usage is not necessarily limited to those who they may consider “paranoid” or who have an unachievable level of specialized knowledge about security.

Summary of Results As They Relate to Hypotheses

In summary, the results from our matched propensity sampling analysis lends additional support to **H2** and conditional support to **H3** and **H4**. Specifically, the prediction, of **H3** and **H4**, that the direction of the effect of social influence on a potential adopter's likelihood to adopt a security feature will shift at a threshold appears to be true for Login Approvals and is likely true for Login Notifications. For Trusted Contacts, however, it appears that social influence has a positive effect on its adoption, regardless of the level of exposure. Furthermore, the observability of a security feature appears to at least partially moderate the presence and value of this threshold.

Discussion

I analyzed whether and how security tools adoptions diffused through the social networks of 1.5 million people who use Facebook. These results provide large-scale empirical evidence that social influence does affect the adoption of security tools, and can do so both positively and negatively. Moreover, the directionality and magnitude of this effect are dependent on at least three factors.

First, the current level of adoption among a potential adopter's friends affects their own likelihood to adopt the tool. While the magnitude of this effect varied across tools, the presence of the effect was consistent across tools. Specifically, for Login Notifications, people who were *unexposed* to a certain percentage of friends who already used the tool were *more* likely to use the tool than those who *were* exposed to that same percentage of friends who already used the tool—for *all* tested levels of exposure. In other words, social proof had a negative effect—the early adopters of Login Notifications appeared to cast a stigma on the tool. For Login Approvals, the same was true up until the highest levels of exposure at which point the exposed were more likely to adopt the tool than the unexposed. In other words, social proof has a negative effect until a certain *critical threshold* of a potential adopter's friends start using the tool, at which point it starts having a positive effect. Finally, for Trusted Contacts, we saw a different trend: For all levels of exposure, those who were exposed were more likely to use Trusted Contacts than those who were unexposed. Accordingly, social proof had a positive effect even at the low levels of exposure to friends who use Trusted Contacts. While each tool was affected differently by social influence, there was one consistency: the effect of social influence got increasingly positive at higher levels of exposure (that is, the difference in adoption rate between people who were exposed and people who were unexposed to friends got increasingly positive at higher discrete levels of exposure).

Second, the difference in adoption trends across tools suggests that the design of a security tool strongly affects its potential for social diffusion. Specifically, using Rogers' Diffusion of Innovations [87] theory and my own prior qualitative work in Chapters 3 and 4 as a lens, it may be that Trusted Contacts has two advantages, in its potential for social spread, over Login Notifications and Login Approvals. First, its use is more *observable*—whereas use of Login Notifications and Login Approvals is private, enabling Trusted Contacts requires a user to specify three to five friends to help with account recovery. These specified “trusted contacts” are, in turn, notified that they have been entrusted with this role and thus its presence is broadcast. Second, Trusted Contacts is more *socially inclusive*. Whereas Login Notifications and Login Approvals are used to *exclude* others from access and may thus be indicative of distrust [1], Trusted Contacts is social—it *includes friends* in the process of improving one's own security, and may thus be more indicative of trust. Indeed, it allows friends to be accountable for each other's security, which, as suggested in the results of our interview study, is something that many present-day security tools are lacking.

Third, exposure to friends from more diverse social contexts who use a security tool may increase one's likelihood to adopt a security tool. Indeed, for all three security tools, controlling for the number of one's friends who used a security tool, people who had exposure to friends from more diverse social contexts who used a security tool had a higher likelihood of adopting the security tool themselves.

To summarize, it seems that **security tool adoption does depend on social influence, but only positively for tools that are observable, socially compatible and/or widely adopted by many distinct social circles within a potential adopter's social network.**

Leveraging Social Influence to Improve End-User Security

The results from this study, as well as those from Chapters 3 and 4, provide strong empirical evidence that social influences affect security. However, the effect of social influence can manifest in nuanced and sometimes unexpected ways. Social influence, for example, appears to have a negative effect on the adoption of many standard security tools such as two-factor authentication. My working model on how social influence affects security behaviors is dual-faceted: the first facet is that security behaviors can carry a social stigma whereby early adopters of security tools drive away laypeople from wanting to use those same tools, and that this stigma can only be overcome through exposure to more and diverse people who use those tools; and, the second facet is that the design of a security tool affects its potential for social spread.

The social stigma of security can be explained by the psychological concept of an *illusory correlation* [20]—i.e., the idea that the attributes of the users of a product can be mistakenly associated as attributes of the product itself. The early adopters of security behaviors tend to be those who are already security experts or are non-experts who may be perceived as paranoid or having something

to hide. Accordingly, many non-experts may believe that strong security behaviors are meant for either experts or those who are paranoid, and if they do not self-identify as an expert or do not want to be seen as “paranoid” they will, in turn, avoid these security behaviors. Supporting evidence for this idea comes from the negative effect of social influence on the adoption of Login Approvals and Login Notifications I presented in this chapter, along with the findings, in Chapter 4, that experts avoid talking about security with non-experts to avoid the perception of being a ‘nag’ or too ‘preachy’. Gaw et al.’s prior work showing that non-experts who encrypt e-mail are viewed as paranoid [45] can also be seen as indicative of this early adopter stigma. Taken together, the idea that social influence works against today’s end-user facing security behaviors neatly explains why, despite years of usability improvements, security sensitivity remains generally low.

The good news, however, is that security doesn’t have to continue carrying this social stigma. My analyses suggest that the design of a security tool affects its potential for social spread. Trusted Contacts, for example, was not negatively affected by social influence, even for those who had just a few friends who used the tool. Furthermore, it appears more generally that the greater the exposure and diversity of exposure to friends who utilize good security tools and behaviors, the more positive is the effect of social influence. Thus, it is possible that we can make the effect of social influence on security behaviors a positive one if we make the behaviors more social. My work points to at least three design dimensions that constitute this sociality: *observability*, *inclusivity* and *stewardship*.

Observability is the idea that good security behaviors should be easy to observe and emulate. Today, security is invisible—it is difficult for people to know who around them cares about cybersecurity and what other people do to ensure their own security.

Inclusivity is the idea that good security behaviors should bring together, not isolate, one from friends and loved ones. Existing security tools are meant for individual use and assume global distrust—a password is one’s own and should never be shared, for example. Perhaps it is no surprise that use of security technologies that assume global distrust carry negative connotations of paranoia.

Stewardship is the idea that there should be some method for people to act on their concern for the security of their friends and loved ones. This could be in the form of an expert sharing good security behaviors with non-experts, or it could be in the form of collective social sensemaking so that even if Alice does not care about her own security, she might be able to help Bob navigate his cybersecurity decisions. Existing security tools afford no outlet for this outward propagation of knowledge—everyone is cloistered in their own digital bubbles.

I propose two prescriptions. The first is to supplement existing security and privacy tools with **socially-inspired interface nudges** that increase their observability, inclusivity and stewardship. In their popular book, *Nudge*, Thaler and Sunstein define a nudge as “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid.” [103] Nudges are not a new concept in usable privacy and security. Acquisti et al. [2] argue that every design choice, in fact, is a nudge and that nudges can help people making good security and privacy decisions when doing so is becoming increasingly difficult. But what constitutes an effective nudge? I argue that socially-inspired nudges that make existing security systems more observable, inclusive or stewarded can help counteract the social stigma of security behaviors. In Chapter 6, I present an experimental evaluation of one such socially-inspired nudge.

The second prescription is to **redesign end-user interactive security systems to be more social**. Socially-inspired nudges may be useful, but are not universally applicable. It would be difficult to make many existing security systems, like Facebook’s Login Approvals, more inclusive or stewarded, for example. I argue, however, that we start re-inventing end-user interactive security systems to be more social by design—by creating systems that are more observable, inclusive and/or stewarded. Doing so will likely require radical shifts in how we think about security and require substantial iteration. But if we are to design systems that encourage better cybersecurity behaviors, it is clear that we need a departure from the vanilla usability-security spectrum. Usability improvements alone appear to be insufficient to raise security sensitivity, and increased security is useful only if the tools and behaviors required are actually used. Instead, we should begin to think about usability, security

and sociality as three independent dimensions of the design space for interactive cybersecurity systems. In Chapter 7, I present an example socially-inclusive authentication system I developed to begin exploring this design space of usable, socially-intelligent security systems.

Chapter 6: Increasing Security Sensitivity with Social Proof

The contents of this chapter are partially drawn from a previously published paper: [Increasing Security Sensitivity With Social Proof: A Large-Scale Experimental Confirmation](#). Sauvik Das, Adam Kramer, Laura Dabbish and Jason Hong. In Proceedings of the 21st ACM Conference on Computer and Communications Security (CCS'14) [29].

Summary

In the previous chapter, I synthesized key insights from Chapters 3 – 5 into a working model of how social influences affect security behavior and put forth two prescriptions: (i) creating socially-inspired interface nudges that add to the observability, inclusivity and stewardship of existing security systems, and (ii) creating novel end-user facing security systems that emphasize these dimensions by design. As an example of the first sort of prescription, I implemented and designed a set of social nudges that increased the *observability* of security behaviors by informing people of their friends' usage of optional security tools on Facebook. To evaluate if these nudges increased the awareness and adoption of the promoted security tools, I ran an experiment. Specifically, I showed 50,000 people who use Facebook one of eight announcements promoting the use of the same three security tools we studied in the previous chapter—Login Approvals, Login Notifications and Trusted Contacts. Seven of the announcements had a social proof cue: i.e., some descriptive text that informed viewers about the fact that some number of their direct friends already used the security tools we were promoting. These social proof cues varied in their *specificity* (i.e., the exact number of friends to just 'some' friends) and *framing* (e.g., "Over X friends" vs. "Only X friends"). The eighth announcement was a non-social control. The results were unambiguous: the social announcements all significantly out-performed the non-social control, increasing clicks on the announcement by 37% and thereby also increasing the number of security tool adoptions by 30%.

Motivation

Looping back to the introduction of this proposal, one of the largest problems in computer security is the need for higher awareness and use of available security tools. From Chapters 3, 4, and 5, I demonstrated that social influence plays a key role in people's awareness and use of available security tools, and identified that security tools and behaviors that were more *observable* might more easily spread via social channels. This finding is line with prior work in social psychology and sociology. Indeed, Rogers, in his seminal Diffusion of Innovations [87], lists observability as one of the five key factors in determining whether a technology will see widespread use. Similarly, studies by Milgram and colleagues [76] and Cialdini and colleagues [21,23,48] also highlight the concept of social proof—or the tendency for one to try and emulate what she can see others around her doing.

Historically, however, security feature usage has been kept confidential to preserve an individual security-tool-user's privacy. While this privacy is important, as using a security tool can also have negative connotations such as being "paranoid" [45] or "nutty" [Chapter 3: A Typology of How Social Influence Affects Security Behaviors, this hiding of security feature use has both stifled the social diffusion of security features *and* made it difficult to test the effect of social interventions on increasing people's security sensitivity. Furthermore, it has been difficult if not impossible to answer these questions because of the lack of data associating security tool adoptions with social meta-data. Consequently, the security community has overlooked a potentially fruitful avenue for increasing security sensitivity, as there is a dearth of empirical data conclusively linking social-proof based interventions to heightened security sensitivity. Today, with the rich and high-complete social meta-data on platforms such as Facebook, we can design simple social cues that show non-adopters social proof that their friends use security tools.

In the following chapter, I share some of the first results experimentally testing whether increasing the observability of security tool usage does indeed increase the awareness and adoption of security tools [3]. Along with colleagues, I designed a set of 7 security announcements with social proof cues that can preserve the privacy of individuals who use security tools while still providing their friends with positive social proof in favor of using the tools. All social announcements informed viewers that their friends used "additional" security tools, but the seven variations differed in their *specificity* (i.e., showing viewers exactly how many of their friends used security features versus just saying that

“some” of their friends used security features) and *framing* (i.e., using keywords such as “only” or “over” to prime viewers’ interpretation of the text). To run this experiment in an ecologically valid setting, I again teamed up with Facebook’s Site Integrity team to promote the security tools we studied in Chapter 5: How Social Influences Affect Security Tool Diffusion: Login Notifications, Login Approvals and Trusted Contacts. The specific research questions we wanted to answer in this experiment were:

Q1. Does increasing the observability of security tool usage drive the awareness and adoption of security tools?

Q2. Does the framing of social information affect the exploration and adoption of security tools? If so, which framings work – those that suggest that many of a friends have already started using security tools, or those that suggest that few have already started using security tools and that the viewer should be among the first?

Q3. Does the specificity of the social proof cue matter? In other words, is it enough to inform viewers that “some friends” use security tools, or is it only effective if they see a specific number?

We ran two experiments. In our first experiment, we tested all of the eight announcements we designed to test whether and which social proof cues yielded the highest click-through rates and follow-up adoptions. In our second experiment, we re-ran only our best performing social conditions and also asked participants to answer a short survey to test whether providing social proof cues in an announcement influenced people’s perceptions of the security tools we promoted—namely, whether a viewer believed the tools were sufficient to address their security concerns.

Experiment

In our initial experiment, we showed 50,000 people who use Facebook one of eight announcements, pinned at the top of their Facebook newsfeed, informing them about the availability of extra security features on Facebook. Seven of these announcements included a social cue informing viewers that their friends also used security features, but varied in their *specificity* (i.e., showing the exact number of friends versus just saying “some” friends) and *framing* (i.e., priming the interpretation of the social cue with keywords such as “only” and “over”). None of the announcements revealed any information about individual tool users, however, thus providing aggregated social proof *without* surfacing *who* was using *which* tools. We measured whether the nature of the text in the announcement (social vs. non-social, the framing and specificity of the social proof text) led to greater exploration of available security features and greater adoption of security features—or, increased *awareness of* and *motivation to use* security features, respectively.

Methodology

People in our sample who logged on to Facebook between November 4th, 2013 and November 8th, 2013 were shown one of eight announcements informing them that they could use extra security features to protect their Facebook accounts. The announcements were rendered at the top of their newsfeeds—the portion of Facebook’s user interface where people are directed when they first log in, where they see an assortment of content shared by their friends (see Figure 8). All announcements contained a call-to-action button (labeled “Improve Account Security”) that directly linked people who clicked on the button to an interstitial that explained the benefits of the three security features we promoted (described below) and allowed viewers to enable the features.

Announcements were shown at most three times to the same person over the course of the four days, in order to mitigate the effect of greater exposure to those who were more active.

Experiment Groups

We designed and implemented four social framings to test not only whether and how social-proof cues can increase people’s security sensitivity, but also if the specificity and framing of those cues matter. We refer to these framings as “Over”, “Only”, “Raw”, and “Some”. The “Over” framing

informed viewers that more than a certain number or percent of their friends use extra security features, priming viewers to interpret the social cue as there being abundant social proof that others they know use security features: i.e., “many people do this, so I should too.” The “Only” framing takes

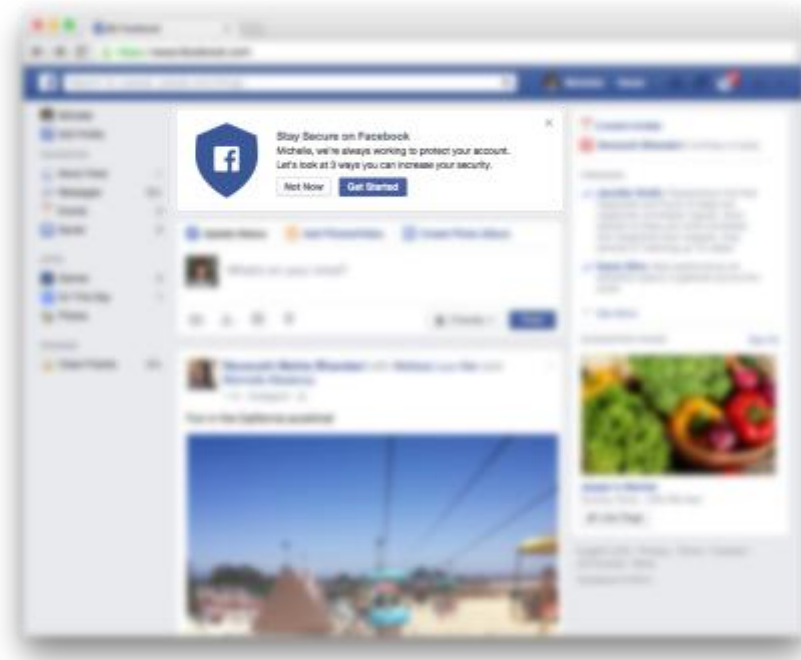


Figure 8. Users saw our announcements on top of their newsfeeds, as shown above, up to 3 times.

a contrasting approach, framing the social cue in a manner that suggests that only a few of a viewer’s friends use security features so they should be among the first of their friends to secure their account. The “Raw” framing eliminates subjectivity in the framing altogether and simply presents the viewer with the quantity of her friends who use security features. Finally, the “Some” framing is ambiguous: informing viewers only that a positive number of their friends use security features.

The “Over”, “Only”, and “Raw” framing had two forms: a *number* form where the number of the viewer’s security-feature using friends was rendered in the announcement, and a *percentage* form where the percentage of the viewer’s security-feature using friends was rendered in the announcement. In total, thus, there was one control group, two “Over” framing groups, two “Only” framing groups, two “Raw” framing groups, and one “Some” framing group, for a total of $1+2+2+2+1=8$ experimental groups. The eight experimental groups are summarized in Table 13, and an image of each of these announcements is shown in Figure 9.

Sample

We selected a random sample of $n=50,000$ people from the U.S. who used Facebook in English, were at least 18 years old, logged on to Facebook at least once in the month preceding the experiment, had at least 10 friends who enabled one of the promoted security features, and had not enabled any one of the security features we were promoting. We evenly assigned the $n=50,000$ people in our sample into one of the aforementioned eight experiment groups, amounting to 6,250 people per group. This assignment was mostly random, with the constraint that people assigned to the Over condition had to have at least 10% of their friends who enabled security features, and people assigned to the Only conditions had to have fewer than 10% of their friends who enabled security features. Our participants were 40 years old on average (s.d., 16), and 68% were women, suggesting that our sampling criteria had a bias towards older females. Notably, our sampling criteria was also biased towards active, non-security experts, but we do not believe this to be a stifling limitation given that

active, non-security experts are the intended target for interventions aiming to heighten security sensitivity, as these people potentially face the greatest risk of having their accounts compromised.

Finally, the $n=50,000$ sample size we selected for our experiment comfortably exceeded the 4,000 participant sample size suggested by a power analysis for generalized linear models [25], with 26 coefficients, a significance level of 0.001, a power of 0.999, and a very modest effect size of 0.02—i.e., a prediction that the best social announcement will only introduce 2% more clicks relative to the control condition. In practice, we expected the effect size to be greater than 2%, but we selected a low effect size for the power analysis to get an upper bound on the number of users we needed to obtain significant results for our experiment. The 26 coefficients in our model comprised of the 18

<i>Group</i>	<i>Prompt Text</i>
<i>Control</i>	You can use security settings to protect your account and make sure it can be recovered if you ever lose access.
<i>Over (#/%)</i>	Over X of your friends use extra security settings. You can also protect your account and make sure it can be recovered if you ever lose access. [Note: X rounded down to nearest 5 th (e.g., 108 becomes 105)]
<i>Only (#/%)</i>	Only X of your friends use extra security settings. Be among the first to protect your account and make sure it can be recovered if you ever lose access.
<i>Raw (#/%)</i>	X of your friends use extra security settings. You can also protect your account and make sure it can be recovered if you ever lose access.
<i>Some</i>	Some of your friends use extra security settings. You can also protect your account and make sure it can be recovered if you ever lose access.

Table 13. Prompt text in announcement across all 8 experimental groups. Some social groups have templates that are filled in with either the number or percentage of a user’s security feature-using friends.

variables described in Table 14, in addition to seven categorical variables representing the experimental conditions, and one intercept variable.

Promoted Security Features

We decided to promote the following three security features in our initial campaign:

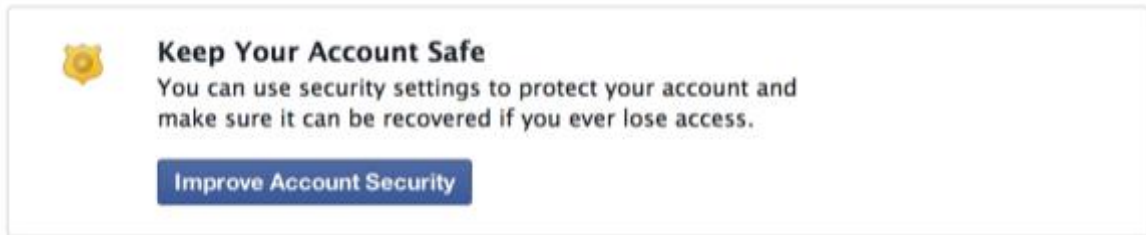
Login Notifications: A security feature that informs users, via text and/or e-mail, whenever their Facebook account is accessed under suspicious circumstances: e.g., from a city the person had not previously visited.

Login Approvals: A two-factor authentication security feature that requires users to enter a randomly generated security code (sent to or generated on their phone) in addition to their passwords in order to authenticate.

Trusted Contacts: A security feature that allows users to specify 3-5 friends who can vouch for the user’s identity if she forgets her Facebook account password and cannot access her e-mail.

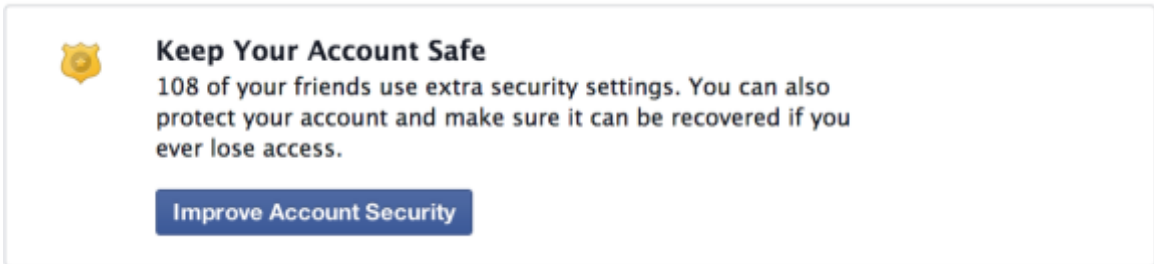
We selected these three security tools because they were the same set of tools we studied in the previous chapter, as well as because they were all co-located within the “security settings” menu context in Facebook’s user interface. We chose to promote three security tools to avoid drawing conclusions specific to any single security tool, and because these tools represented a wide range of definitions for “security features”—with Login Notifications simply informing people of potential breaches, Login Approvals adding an extra step to the authentication process, and Trusted Contacts asking people to draw in their friends to help protect their accounts.

1. Control



This screenshot shows a control announcement. On the left is a yellow shield icon. To its right, the text reads: **Keep Your Account Safe** followed by "You can use security settings to protect your account and make sure it can be recovered if you ever lose access." Below this text is a blue button with the text "Improve Account Security".

2. Raw #



This screenshot shows an announcement with the framing "Raw #". It features a yellow shield icon, the text **Keep Your Account Safe**, and the message "108 of your friends use extra security settings. You can also protect your account and make sure it can be recovered if you ever lose access." A blue button labeled "Improve Account Security" is positioned at the bottom.

3. Over %



This screenshot shows an announcement with the framing "Over %". It features a yellow shield icon, the text **Keep Your Account Safe**, and the message "Over 20% of your friends use extra security settings. You can also protect your account and make sure it can be recovered if you ever lose access." A blue button labeled "Improve Account Security" is positioned at the bottom.

4. Some



This screenshot shows an announcement with the framing "Some". It features a yellow shield icon, the text **Keep Your Account Safe**, and the message "Some of your friends are using extra security settings. You can also protect your account and make sure it can be recovered if you ever lose access." A blue button labeled "Improve Account Security" is positioned at the bottom.

5. Only #



This screenshot shows an announcement with the framing "Only #". It features a yellow shield icon, the text **Keep Your Account Safe**, and the message "Only 108 of your friends use extra security settings. Be among the first to protect your account and make sure it can be recovered if you ever lose access." A blue button labeled "Improve Account Security" is positioned at the bottom.

Figure 9. Screenshots of each of the social framings and the control announcements in our experiment. In total we had 8 announcements. Not pictured is the Raw %, Over # and Only % announcements, but they look similar to their counterparts pictured above.

Dataset

We measured click-through rate for each announcement, as well as the short-term and long-term adoption rate of the promoted security features up to a week and 5 months after running the experiment, respectively. We used click-through rate on the announcement as a proxy for raising *awareness* (as people who clicked on the announcement were taken to explore the promoted security features), and adoption rate as a proxy for raising *motivation* (as people who adopted security features must have gained the motivation to enact a behavior change). We could not measure the differential effects of the announcements on *knowledge*, however, as all announcements led viewers to the same interstitial with the same information.

In addition, we collected each viewer’s number of friends who used any of the three security tools we were promoting, along with a set of behavioral (e.g., frequency of posts and comments), demographic (e.g., age, gender) and social network descriptor (e.g., mean friend age, mean friend-of-friend count) control variables that we expected might affect click-through rate and security feature adoption

Demographic Variables	
<i>Age</i>	Age of the user.
<i>Gender</i>	Self-reported gender: male or female.
<i>Friend count</i>	Count of the user’s number of friends.
<i>Account length</i>	Days that have passed since the user activated his/her account.
Social Network Variables	
<i>Mean friend age</i>	Average age of friends.
<i>Friend age entropy</i>	Shannon entropy of friend ages.
<i>Percent male friends</i>	Percentage of friends that are male.
<i>Mean friends’ account length</i>	Average number of days the user’s friends have used Facebook.
<i>Friend country entropy</i>	Shannon entropy of countries from which the user has friends.
<i>Mean friend of friend count</i>	Average number of friends of friends.
Behavioral Variables (all aggregated across the week prior to data collection)	
<i>Posts Created</i>	Number of posts created.
<i>Posts Deleted</i>	Number of posts deleted.
<i>Comments Created</i>	Number of comments created.
<i>Comments Deleted</i>	Number of comments deleted.
<i>Friends Added</i>	Number of friends added.
<i>Friends Removed</i>	Number of friends removed.
<i>Photos Added</i>	Number of photos added.
Social Variables	
<i>Feature-using friends</i>	Number of friends who use security features.

Table 14. Collected feature descriptions and distributions for the $n=50,000$ people in our sample. † Approximate values.

among our sample. These variables are described in Table 14.

Hypotheses

Cialdini's [21] concept of *social proof* suggests that when we are confronted with making a decision where we are uncertain of the appropriate course of action—like adopting a security tool, say—we look to our friends and those around us for cues on how to act. Combined with Rogers [86] assertion that *observability*—or, the visibility of the use and benefits of an innovation—is critical to the widespread adoption of an innovation, our own finding that the observability of security tool usage is a major positive factor in security and privacy related behavior change, we predicted:

H1: Social announcements will have higher click-through rates than the non-social control.

Extending the idea that social proof is more convincing when people see larger groups conforming to an action [76], we also predicted:

H2a: People with more security-tool using friends will be more likely to click on the announcement.

H2b: People with more security-tool using friends will be more likely to adopt a security tool, both in the short and long-term.

Similarly, we predicted experiment groups that rendered higher values or otherwise suggested that *more* rather than *fewer* of the viewer's friends used security features would be more effective at getting users to click on the announcement and explore security features. Thus, we expected that "number" conditions would have higher click-through rates than their "percent" counterparts, as the former generally render higher numbers in the announcement (e.g., **20** friends vs. $20/400=5\%$ of friends). Furthermore, as the "Raw" framing rendered the highest values, followed by the "Over" and then the "Only" framing, we expected that the click-through rates for these framings would fall in that order as well.

H3a: The "number (#)" context conditions will have higher click-through rates than their "percent (%)" counterparts.

H3b: The "Raw" framing will have the highest click-through rate, followed by the "Over" and then "Only" framings.

Next, as one of the driving forces for social proof is a search for a clear course of action in an unclear circumstance [21], we also suspected that clearer, more informative messages would be more effective at driving click-through rate.

H4: More specific social framings will have higher click-through rates. Thus, the "Some" context will have the lowest click-through rate.

For short-term adoptions, we expected that the effects of social conditions would be muted. Indeed, while it is cheap—in terms of time and effort—for people to explore and gather information about security features, it can be expensive for them to actually activate those features. For example, activating Login Approvals would require people to spend an extra few seconds every time they "logged in" to their Facebook accounts. Taken together with the previous finding that people generally only enact security and privacy related behavior change after personally experiencing or hearing about a threat [28], and Egelman and colleagues' finding that a "peer pressure" password meter did not raise people's motivation to create stronger passwords relative to a non-social password meter [40], we expected that, in the short term, there would be no difference in security feature adoption rate among those who view social and non-social announcements.

H5: The adoption rate for the promoted security features should be about the same for those who view a social or a non-social announcement in the week following the experiment.

On the other hand, we expected that there *should* be a long-term increase in the overall security feature adoption among users in the social condition. While our experiment lacked a strong *catalyst* for security behavior change, we expected that people in the social conditions might more strongly retain the information that extra security features are available for when they *do* encounter a compelling catalyst (e.g., hearing about a security breach on the news or through a friend). As a number of highly publicized security vulnerabilities were surfaced in the five months following the experiment (including the widely publicized “Heartbleed” bug in OpenSSL [115]), we arrived at:

H6: The adoption rate for the promoted security features should be higher for those who view a social announcement compared to those who viewed a non-social announcement in the 5 months following the experiment.

Results

Out of the 50,000 people in our sample, 46,235 logged in to Facebook within the duration of the experiment and were shown an announcement. Across all conditions, 5971 (13%) people clicked on the announcement to explore the promoted security features, while 1873 (4%) people adopted one of the promoted security features within the following week, and 4555 (9.9%) within the following five months. In Table 15, we show an aggregated breakdown of clicks and adoptions across experiment groups. The raw data suggests that all social conditions had higher click-through rates than control, the best social announcements elicited higher adoption rates in the short and long term, and the “Raw #” announcement generally performed best of all.

Group	N	Clicked	Adopted	Adopted
All Conditions				
Raw #	5862	846	280	623
Some	5828	835	243	602
Over #	5770	779	248	547
Only #	5668	748	225	548
Over %	5761	724	223	557
Only %	5708	714	221	555
Raw %	5953	730	225	573
Control	5685	595	208	550
Social vs. Non-Social				
Social	40550	4376	1665	4005
Control	5685	595	208	550
Social Number vs. Social Percent				
Number	17300	2373	753	1718
Percent	17422	2168	669	1685
Social Contexts				
Raw	11815	1576	505	1196
Over	11531	1503	471	1104
Only	11376	1462	446	1103

Table 15. Clicks and adoptions by experimental conditions. “N” represents the number of users who viewed the announcement. “ST” stands for short term, and “LT” stands for long term. These values are strictly descriptive. Statistical tests used and significance is mentioned where relevant in the text.

To statistically test whether and how the existence of, specificity, and framing of the social cue in the announcement affected click-through rate and security feature adoption, we ran three logistic regressions for clicks, short-term adoptions, and long-term adoptions. The response variables for our three models were, respectively, binary values representing (i) whether or not an individual had clicked on the announcement they were shown, (ii) whether or not an individual had adopted any of the three promoted security features in the 7 days following our experiment, and (iii) whether or not an individual had adopted any of the three promoted security features in the 5 months following the experiment. Our independent variable was which of the eight social announcement an individual had seen, and we also included, as controls, the behavioral, demographic, and social network descriptor variables listed in Table 14. For the two adoption models, we included an additional control

representing whether or not an individual had actually clicked on the announcement they were shown to “Improve Account Security”.

In Table 16, we show the logistic regression coefficients for our independent variables predicting clicks, short-term adoptions and long-term adoptions. Coefficients in Table 16 represent a change in “log-odds”, or $\ln \frac{P}{1-P}$, where P represents the probability that the user clicked on the announcement or adopted one of the three security features, depending on the model. A positive coefficient implies that the log-odds ratio increases, or that the variable for the coefficient increases the likelihood that the viewer clicked on the announcement or adopted a security feature. A negative coefficient implies the opposite. Furthermore, all variables are centered and scaled, such that the coefficient for each variable represents the expected change in log-odds that an individual uses a feature given a one standard deviation increase in the predictor variable, holding all other numeric variables at their means and categorical variables at their baselines. Additionally, larger absolute coefficient values imply a stronger relationship between the independent and dependent variables.

Variable Name	Clicked	Adopted 7-day	Adopted 5-month
† Group: At Least #	0.29 *	-0.07 *	-0.13
† Group: At Least %	0.21 *	-0.12	-0.06
† Group: Only #	0.26 *	-0.16	-0.09
† Group: Only %	0.19 *	-0.12	-0.05
† Group: Raw #	0.36 *	-0.01	-0.001
† Group: Raw %	0.17 *	-0.15	-0.06
† Group: Some	0.35 *	-0.18	-0.03
Tool-using friends	0.09 *	0.17 *	0.20 *
Intercept	-2.16 *	-5.23 *	-2.62 *
Age	-0.01	-0.19 *	-0.18 *
Gender: Male	-0.03	-0.06	-0.13 *
Account length	0.11 *	0.03	0.03
Friend count	-0.16 *	-0.06	-0.15 *
Mean friend age	0.14 *	-0.16	-0.24 *
Friend age entropy	0.03	0.28 *	0.26 *
Percent male	0.02	0.08	0.13 *
Mean friends days since confirmed	0.007	0.003	-0.08 *
Friend country entropy	0.04 *	0.04	0.03
Mean number of friends of friends	-0.04	-0.09	-0.09 *
Posts created	0.05	0.02	-0.02
Posts deleted	-0.008	0.02	-0.002
Comments created	0.09 *	0.07 *	0.10 *
Comments deleted	0.07	-0.13	-0.01
Friends added	-0.003	0.004	0.02
Friends removed	-0.004	0.01	0.02
Photos added	0.03 *	0.004	0.03
Clicked on Announcement	NA	4.38 *	1.94 *

† Baseline: Control, * p < 0.05

Table 16. Coefficients for the three regressions predicting clicks, feature adoptions up to a week after the experiment, and feature adoption up to 5 months after the experiment. Bolded coefficients are of interest..

For example, the *tool-using friends* variable (i.e., the number of one’s friends who use security features) coefficient for the “clicks” model is 0.09; thus, a one standard deviation increase in this variable increases the log-odds that a viewer clicks on the announcement by 0.09, and the actual odds by $e^{0.09} = 1.09$. More concretely, our model predicts that someone with 80 security feature-using friends (one standard deviation above the mean) is 9% more likely to have clicked on the security announcement, compared to the average person in our sample.

From Table 16, we can see that, relative to the control condition, all social experiment conditions do elicit higher click-through rates for announcements, as evidenced by the positive *and* significant coefficients for every experiment condition coefficient. The “Raw #” ($b_{clicked}=0.36$, $p<0.001$) condition had the highest click through rate, at 14.4%—a substantial 37% increase relative to control. Even the

least effective social condition—the “Raw %” ($b_{clicked}=0.17$, $p<0.001$) condition—significantly enhanced click-through rate relative to control, up to 12.3%. There does, therefore, appear to be strong evidence in favor of **H1**—that all social conditions will improve click-through rate relative to the control condition. The effect is both significant and substantial.

There is also support for both **H2a** and **H2b**—that people with more security-tool using friends will be more likely to click on the announcement and adopt a promoted security tool. The *tool-using friends* variable ($b_{clicked}=0.09$, $p<0.05$; $b_{adoptions-7d}=0.17$, $p<0.05$; $b_{adoptions-5mo}=0.20$, $p<0.001$) has a large and positive coefficient for all three models, suggesting that viewers who see that more of their friends use optional security tools are more likely to click on the announcement *and* actually adopt a security tool relative to the average person in our sample (with all numeric variables at the mean and categorical variables at the baseline).

The data, however, is not as clear in its support for the hypothesis that social framings that suggest *more* rather than *fewer* of a viewer’s friends use security tools will be more effective at driving click-through rate on the security announcement. There does appear to be support for **H3a**—that number conditions will outperform percent conditions in driving click-through rate on the security announcement. Indeed, all number conditions significantly outperformed all percent conditions, and, in aggregate, number conditions elicited 7% more clicks than percent conditions ($\chi^2(1, n=34,722)=12.3$, $p=0.0004$). However, we found no support for **H3b**—that the “Raw” framing would outperform the “Over” framing, which, in turn, would outperform the “Only” framing in driving click-through rate. While the aggregated click-through rate of these framings do fall into the expected sequence (Raw=13.3%, Over=13.0%, Only=12.9%), the difference is not significant despite massive power ($\chi^2(2, n=34,722)=1.2$, $p=0.54$).

Thus, while social announcements that suggest that *more* rather than *fewer* of a viewer’s friends are currently using extra security features can be more effective at getting people to click on the announcement, the specific framing of the social text does not appear to significantly impact its click-through rate.

Relatedly, we found evidence to contradict **H4**—that ambiguous social framings such as the “Some” framing will be less effective at driving click-through rate for the announcement. In fact, the “Some” ($b_{clicked}=0.35$, $p<0.001$) framing is the second most effective group in driving click-through rate, after the “Raw #” ($b_{clicked}=0.37$, $p<0.001$) condition, with an overall click-through rate of 14.5%.

We derived **H4** from a simple understanding of social proof—if people look to their friends for cues on how to act during periods of uncertainty, then ambiguous cues are probably less effective than clear cues. However, in reality, the ambiguity appears to elicit more interest in the announcement than most of the more specific social framings. Perhaps this finding can be explained by the intuition that people may overestimate the number of their friends who use security features when it is left ambiguous. Future work can validate this hypothesis by looking at the discrepancy between people’s perceptions of the number of their friends who use security tools relative to the actual number of their friends who use security tools.

Next, there appears to be support for **H5**—that social prompts will not be significantly more effective at driving feature adoption in the short-term than non-social prompts. Indeed, all of the coefficients for the social conditions are insignificant in the short-term adoptions model in Table 16. We expected this result for two reasons: (1) people usually only adopt security tools after experiencing a “catalyst” for security behavior change—for example, in the form of experiencing a security breach or hearing about a security breach [28], and (2) the social text is not reinforced in the security interstitial where people must actually make the decision to adopt a security feature—thus, as with Egelman and colleagues’ study [40], potential adopters are *not* given enough social context at the moment of potential behavior change—for example, who among their friends use what security tools.

More surprising, however, is that this negative result holds even for long-term adoptions, disconfirming **H6**—that social announcements *will* be significantly more effective at driving security feature adoption in the long term relative to the non-social announcement. In the 5 months following the experiment, a number of widely publicized security vulnerabilities that could have served as

catalysts for security behavior change were highly publicized (e.g., Heartbleed [115], the iOS SSL bug [116]). Nevertheless, there was no significant difference in adoption rate between those who saw the social and non-social announcements, perhaps because the social announcements were not more memorable. We also note, however, that H6 may in fact be valid, but only with respect to relevant security threats that are presented on time and in context: Activating Login Approvals would not have been a direct answer to Heartbleed or the iOS SSL bug, so the latter may not have easily triggered a memory of the former.

Importantly, the immediate *cascading* effects of raising people's awareness of security features should not be ignored. While there is no significant difference in the *rate* of feature adoption between people who *clicked on* either the social or non-social announcement, as significantly more people clicked on the social announcements, many more people who saw social announcements also actually *adopted* security features. Indeed, from Table 15, we can see that 280 of 5862 (4.8%) people shown the "Raw #" announcement adopted one of the promoted security features over the 7 days following the experiment, compared to just 208 of 5685 (3.7%) people shown the non-social announcement ($\chi^2(1, n=11,547)=8.7, p=0.003$). In other words, significantly more people who saw a social announcement adopted the promoted security features because significantly more people *clicked on* the social announcements.

Summary of Results from Initial Experiment

We found that increasing the observability of security tool usage can be effectively used to increase both *awareness* of and *adoption* of available security features. Furthermore, this effect increases with the number of the viewer's friends who use security tools. While neither the framing of a social cue nor its specificity appeared to have a large effect on raising click-through rate, social announcements that rendered the number of a viewer's friends that used security tools, rather than the percent of the same, elicited higher click-through rates. On the other hand, we found no evidence that the social proof cues we tested, which were aggregated and anonymous, were more effective than a non-social announcement at raising a viewer's *motivation* to use the promoted security features. Indeed, the rate of feature adoption among viewers who clicked on any of the announcements were non-significantly different despite massive statistical power.

Follow-up Study With Survey to Gauge Sentiment and Awareness

To more concretely measure whether our announcements increased people's awareness of available security features, we ran a second deployment of our best performing announcements from the initial experiment and collected survey responses.

Methodology

We re-ran a second campaign of our experiment with a separate set of $n=50,000$ people, randomly sampled among across users who used Facebook in English, logged in to Facebook at least once in the past month, and had at least 10 friends who used security features. People in our sample were shown one of three announcements mirroring the announcements in the previous experiment: the unambiguous "raw number" social condition, the ambiguous "some" social condition, and the non-social control condition—all exactly matching the corresponding condition from the initial experiment. All announcements were once again outfit with an "Improve Account Security" button that, when clicked, would navigate the clicker to an interstitial that explained the promoted security tools, as well as allowed viewers to enable the same. The follow-up study ran between December 20th and December 22nd, 2013.

In this second campaign, we also asked people to complete a short survey with the following 3-point Likert-scale question: Facebook provides me with the necessary security settings to protect my account (i.e., the "Provides security tools" statement). We decided to ask this question to test whether social information in the announcement influenced people's perceptions of the security tools we promoted—namely, whether a viewer believed the tools were sufficient to address their security concerns.

We had three methods to solicit survey responses. First, we surveyed people who fully navigated through the interstitial (i.e., the "interstitial" solicitation group). We separately sent the survey to

people who saw an announcement but never clicked on it (i.e., the “viewed announcement” solicitation group), and also to a random sample of 80,000 people who used Facebook in English, logged in to Facebook at least once in the past month, and who never viewed any of our security announcements (i.e., the “holdout” solicitation group).

In total, we had 2814 responses to our survey. Table 17 shows a tabulation of the how many users per experimental condition and survey solicitation method.

	<i>Holdout</i>	<i>Non-Social</i>	<i>Raw #</i>	<i>Some</i>
<i>Interstitial</i>	0	498	226	254
<i>Viewed Announcement</i>	0	127	72	67
<i>Holdout</i>	788	322	214	246

Table 17. Number of survey responses per solicitation method (rows) and experimental group (columns)

Results

Table 18 shows the coefficients for a proportional-odds logistic regression [52] predicting the likelihood of an individual selecting a higher value of agreement with the “Provides security tools” statement previously explained. Coefficients in Table 18 represent a change in “log-odds” that the user selected “neutral” over “disagree” or “agree” over “neutral” as a response to one of the questions. We included the viewer’s experiment group as well how they were solicited to complete the survey as independent variables, and included the behavioral, demographic and social network descriptor variables described in Table 14 as controls.

	<i>Variable name</i>	<i>Provides security tools</i>
†	Group: Non-Social	-0.08
†	Group: Raw #	-0.19
†	Group: Some	-0.16
Δ	Solicitation: Interstitial	1.04 *
Δ	Solicitation: Viewed Announcement	0.16
	Feature-using friends	-0.13
	Age	0.04
	Gender: Male	0.15
	Account length	0.20 *
	Friend count	0.25 *
	Mean friend age	-0.14
	Friend age entropy	0.07
	Percent male	0.03
	Mean friends days since confirmed	-0.57 *
	Friend country entropy	0.005
	Mean number of friends of friends	-0.08
	Posts created	0.02
	Posts deleted	-0.07
	Comments created	-0.06
	Comments deleted	0.05
	Friends added	0.05
	Friends removed	-0.05
	Photos added	-0.001

† Baseline: Holdout; Δ Baseline: Holdout, * $p < 0.05$

Table 18. Coefficients for the two proportional-odds logistic regressions predicting agreement with the trustworthy and protection statements.

† Baseline: Holdout; Δ Baseline: Holdout, * $p < 0.001$

Just as in the previous study, a positive coefficient implies that the log-odds ratio increases, or that the variable for the coefficient increases the likelihood that the user selected “neutral” over “disagree” or “agree” over “neutral”. A negative coefficient implies the opposite. Furthermore,

predictor variables were centered and scaled, such that each coefficient represents the expected change in log-odds that the user selected a higher value response given a one standard deviation increase in the predictor variable, holding all other numerical variables at their means and categorical variables at their baselines.

From Table 18, there appears to be no significant effect of viewing any of the security announcements on people's agreement with Facebook providing necessary security features, helping to explain why we saw the same adoption *rate* among both those who saw social and non-social announcements. Indeed, none of the coefficients for the "Group" variable were significant.

On the other hand, people who actually clicked on the announcement and navigated through the security interstitial were significantly and substantially more likely to agree with the "Provides security tools" statement ($b=1.04$, $p<0.001$) statement. Thus, while showing people security announcements with social information does not appear to directly affect people's sentiment towards Facebook's security tools, social announcements drive more people to the security interstitial and thus can at least *indirectly* raise their awareness or available security tools and their belief that those security tools are effective.

Discussion

In a nutshell, these results suggest that social proof is a promising approach to increase people's security sensitivity, but it is not a panacea. People who saw announcements with social proof cues that increased the *observability* of security tool usage were more likely to click on the announcement. Clicking on this announcement, in turn, increased viewers' (i) *awareness* of available security tools, (ii) their likelihood to *adopt* one of those tools, and (iii) their *sentiment* towards the efficacy of the promoted tools. However, the aggregated, impersonal social information I showed people only seemed to raise their interest in *exploring* security tools—I did not find strong evidence that the social proof cues, themselves, were more effective than a non-social announcement in increasing people's likelihood of actually adopting one of the promoted security tools (though the results do not prove the opposite, either).

Aggregate social proof cues can raise people's interest in exploring available security tools.

The positive effect of these social announcements on click-through rate is especially strong when viewers have many friends who use security tools and when that information is rendered directly in the announcement, as with the "Raw #" announcement—a finding aligning with both the concept of *social proof* [21] and the *diffusion of innovations* [86]. This result suggests that the positive effect of these social cues will strengthen over time as more and more people start using security tools (and thus higher and higher numbers will be rendered in the announcement). We also found evidence that social announcements *indirectly* appeared to increase viewers' belief that the security tool they needed to secure their accounts were available. Indeed, people who viewed a social announcement were far more likely to click on the announcement and navigate through the resulting security interstitial, and people who navigated through the security interstitial were far more likely to agree that Facebook provided them with necessary security tools.

However, these social proof cues, alone, were not more effective at getting people who clicked on the announcement to actually adopt a promoted security tool. Thus, used alone, the social announcements we tested appeared to be no better than a non-social announcement at raising users *motivation* to adopt the promoted security features. This finding holds true in both the short and long term, even through a number of widely publicized security vulnerabilities including Heartbleed [115] and the iOS SSL implementation bug [116] that could have been potential catalysts for security behavior change [28]. Nevertheless, as more people who saw a social announcement *clicked* on the announcement and explored the promoted security tools, significantly more people who saw a social announcement *adopted* one of the promoted security features. There was, thus, an indirect increase in security tool uptake as a result of showing people a social announcement.

Social proof cues might be more effective if more personal and shown in context. Importantly, these findings do *not* suggest that social cues are ineffective at raising people's motivation to use security tools. Rather, the null result at raising motivation was likely an artifact of the fact that the prompts I tested were aggregated, out of context and not very informative. For example, showing

someone an announcement that 100 of her friends use security tools does not inform her *why* those friends use security tools, *which* security tools are being used (or for what purpose), *who* among her friends are using those security tools, and whether or not her friends would actually *recommend* using those tools. In other words, our absence of results in raising motivation may be due to lack of compensation for an invalid *context*—i.e., asking people to consider extra security tools when they are not really thinking about security. Accordingly, motivation to adopt security tools might be best driven by a paired approach of security threat detection followed by a timely delivery of a security announcement with social cues.

Taken together, in this experiment, I have provided some experimental evidence that simple social proof cues *can* be used to raise peoples' security sensitivity—specifically, their awareness of available security tools. Furthermore, using these simple social cues may have the additional indirect benefits of raising security tool adoption and people's sentiment towards the promoted tools, as well. Care should be taken, however, to sparingly surface these announcements so that people do not get desensitized to them. For example, to maximize the efficacy of a campaign to raise security sensitivity, social announcements should only be shown occasionally to people who already have many friends who use the security tools promoted in a campaign.

My results in this study suggest that *socially-inspired interface nudges* are a promising mechanism through which to encourage better cybersecurity behaviors. Still, there is a vast design space for these nudges that my work only begins to explore. Accordingly, there may be much room for improvement.

Chapter 7: Thumprint

The contents of this chapter are partially drawn from a previously published paper: [Thumprint: Socially-Inclusive Group Authentication Through Shared Secret Knocks](#). Sauvik Das, Gierad Laput, Chris Harrison and Jason Hong. In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI'17) [31].

Summary

In Chapters 3-5, I constructed behavioral models of how social influences affect security behaviors and tool adoptions. I then synthesized the findings from those empirical investigations into a set of three important but rarely considered social design dimensions for interactive cybersecurity systems—*observability*, *inclusivity*, and *stewardship*. In order to better explore this new design space to create more social, interactive cybersecurity systems, I put forth two prescriptions: (i) creating interface nudges that add to the observability, inclusivity and stewardship of existing security systems, and (ii) creating novel end-user facing security systems that emphasize these dimensions by design. In Chapter 6, I designed, implemented and evaluated an instantiation of the first sort of prescription. In this chapter, I present an example of the second—an *inclusive* authentication system designed for small, local groups (e.g., families, work teams, student organizations).

These small, local groups who share protected resources often have unmet authentication needs. Specifically, for these groups, existing authentication strategies either create unnecessary, hard divisions to access shared group resources (e.g., biometrics), do not identify individuals (e.g., shared passwords), do not equitably distribute security responsibility (e.g., individual passwords), or make it difficult to share and revoke access (e.g., physical keys). To explore an alternative, I designed Thumprint. In brief, Thumprint authenticates groups based on group members' expression of a shared, three-second knock on a surface instrumented with (or containing) an accelerometer and microphone (see Figure 10). As the secret knock is shared, group members need not maintain their own individual secrets. However, because individual expressions of the knock are variable, Thumprint can still identify individuals. Current members can safely share the secret with new members, but as individuals are identifiable, previous members can have their access revoked or limited. Notably, Thumprint is not designed to provide perfect security—it is designed to be lightweight and inclusive.



Figure 10. With Thumprint, groups of users learn a single, shared secret knock that they enter on a surface instrumented with (or containing) an accelerometer and microphone (here, a smartphone) in order to authenticate.

To evaluate the usability and security of Thumprint, I ran two user studies. Through these studies, I found that (1) different people who enter the same thumprint can be recognized, (2) people can consistently enter their thumprints over time-separated sessions, and (3) thumprints are fairly secure against casual adversaries—comparable to existing behavioral biometric techniques such as keystroke dynamics [58] and TapSongs [112], but with an added social component.

Motivation

Authentication is important for any secure system, but is typically designed for individuals with privately owned resources and a strong desire to protect them (e.g., bank statements, emails) [26,94]. This focus, while important, has resulted in authentication systems (e.g., PINs, biometrics) that are often inappropriate for a large spectrum of small, local groups who have relaxed security needs and *collectively* share accounts, devices and/or spaces, for example, families who share tablets with children. Shared passwords and PINs do not allow for parental controls, whereas requiring individual passwords for each family member is unwieldy and often subverted [69,74].

Another example is interest-based organizations that share equipment (e.g., a tennis club). Each group member should have access to this shared equipment, but group members often change so using a shared password or key can make it difficult to revoke access from old members [9]. Conversely, use of individual secrets to access group resources can be socially inappropriate [28,45] or rude [25]. Similar situations arise with, for example, employees who share kitchenettes, waitstaff who share access to employee-only areas, and roommates who share a Netflix account.

Despite the prevalence of these local group units, their authentication needs have scarcely been studied in their own right. Through a survey and synthesis of the existing literature on social psychology and usable security, I identified at least three important considerations outside of outsider rejection.

The first is *inclusivity with identifiability* which is, of course, related to the social design dimension of inclusivity I emphasize in Chapter 6. In Facebook's Trusted Contacts, inclusivity represented the idea of including friends in the process of providing oneself security. In the case of group authentication, I define it as reducing the need for individual secrets when authenticating to common group resources. Indeed, requiring individual secrets (e.g., private passwords) to access shared resources is cumbersome and can lead to non-compliance (e.g., sharing passwords) [74]. Individual secrets can also have social consequences: not sharing these secrets can be rude [104] or otherwise create social friction between people [16,28,66,99].

Consider the case of a spouse needing her partner to check a shared calendar on her smartphone: What should she do if the phone is password protected? If the choice is between losing social capital with a loved one or sharing a password, people often opt for the latter [45,107], and, in so doing, break the security assumptions of the system.

Still, social group structures vary widely [109], and some group structures may require access control at the individual level. Thus, inclusivity should ideally come with identifiability to allow for audit logs, personalized functionalities, and tiered access to resources. For example, having one shared family PIN prevents individual family members from creating personalized profiles and precludes the ability for parents to have privileged access [16,37].

Another need is *proportionate distribution of security responsibility*, which is somewhat related to the social design dimension of stewardship I listed in Chapter 6. Responsibility for the well-being of the group should be appropriately distributed across members [65,72], as individuals may be resistant to weighty security solutions that require a large personal investment of time or effort. Accordingly, authentication should be sensibly simple for individual group members. Otherwise, individuals who are less knowledgeable or motivated about security could compromise the whole group's security (e.g., by creating a weak password to access group resources).

One example of where proportionate security responsibility is employed is when nursing staff must use authentication to access hospital computing systems. Nurses often have urgent needs and cannot each be expected to remember long, complex passwords, even if hospital IT has a different perspective. Doing so sometimes results in nursing staff writing down passwords for sensitive hospital equipment right on the apparatus [69].

Finally, there are a number of small, local groups that are built off of a common-identity (e.g., a common interest in tennis). Typically, groups like these have a lot of churn: i.e., they often gain new members and lose old members [80]. For these groups, it should be *easy to share access with new*

members and revoke access from old members. Student organizations that have members rotating every semester, for example, need a simple and reliable way to revoke and grant access to shared equipment closets.

Few existing security solutions support these needs as core functionality. Indeed, in a longitudinal field study of the access control habits of a local group who shared a work space, Bauer et al. found that existing strategies for authentication and access control (i.e., sharing physical keys) could not support the group's ideal policies [9], which accords with Ackerman's broader argument of the socio-technical gap between the social requirements and the technical capabilities of computing systems [1]. Taken together, it appears that a more nuanced social approach can make authentication more usable and useful.

While there has, so far, been little work on creating better local group authenticators, there has been some promising research that explores the problem domain. Toomim et al. introduced a photo access control mechanism where the correct audience should be able to answer a question based on shared knowledge [104]. Gilbert created a social encryption tool, OpenBook, that obfuscates messages in a way that can only be reconstructed by the shared social context between sender and receiver [46]. And, Egelman et al. and Brush introduced the "Family Account" [16,37]—a shared account for all family members. Still, Family Accounts are for access control, not authentication.

One solution is to create a form of authentication that allows group members to share just a single secret but that can still identify individuals. Sensible gestures and mechanical expressions are one promising direction: the shared secret can be the gesture, while each individual might still be predictably unique in their expression of the gesture. My key idea with Thumprint is to use physical knocks as shared group secrets that have varying individual expressions. For instance, accelerometers in mobile devices have been used for detecting a wide range of gestures, activities and hand postures [47]. There has been increasing interest in using these forms of sensible user behavior for authentication. One notable example is the use of keystroke dynamics (i.e., the rhythm with which people type) for authentication [58,77]. With TapSongs, Wobbrock extended this approach to intentional behaviors in the form of rhythmic up-down taps on a binary sensor to match a known jingle timing model [112]. Lin, Ashbrook and White used a similar approach to pair I/O constrained devices through entry of a secret "tapword" on both devices [73]. In all of these cases, outsider rejection was not perfect (~20% failure rates), but insider acceptance was promising.

These approaches, while inspirational, were not designed to be inclusive nor were they meant for groups. With Thumprint, I extend these advances in sensing intentional behaviors for group authentication in order to begin exploring the space of "inclusive" cybersecurity.

System Design

Design Inspiration

There are many analogues in the offline world that illustrate the use of shared secrets for group authentication [11]. There is the famous biblical example of correctly pronouncing the word "shibboleth" that the Gileadites used to identify the invading Ephraimites who could not pronounce the "sh" sound [117]. Other examples include secret handshakes (e.g., the use of selective pressure in handshakes) and code phrases (e.g., saying the words "open sesame" to gain access to a secret lair) [11]. In all of these cases, the shared secret not only authenticates, but is inclusive and reinforces group cohesion [107]. Thumprint is inspired by the secret knocks used at speakeasies during the Prohibition era of the U.S. [67]. At that time, secret knocks were used to identify prospective bar patrons when sale of alcohol was illegal. They could only be learned through social channels and knowledge of a secret knock identified an unknown stranger as part of a broader social collective.

Borrowing from that pattern, Thumprint authenticates local groups with a secret knock consisting of a shared secret token and pattern. The token can simply be a finger or a knuckle, but any small, solid object can be used (e.g., a pen or coin). The pattern can be any sequence of knocks within a three-second period. Authentication occurs by entering the knock on a sensor surface. Furthermore, as each person mechanically enters the knock differently, Thumprint can also identify individuals.

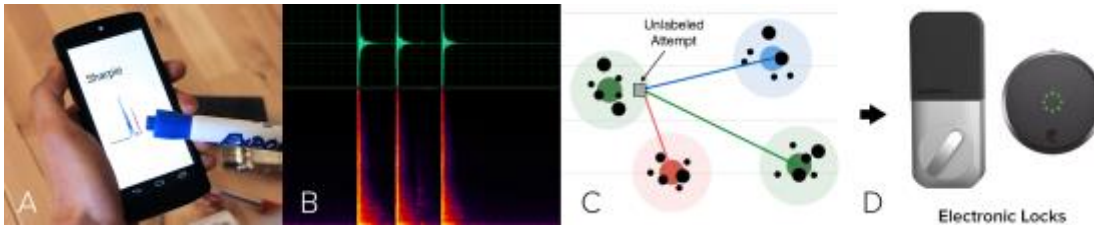


Figure 11. With Thumprint, users enter secret knocks on an instrumented sensor surface (A) from which a variety of time and frequency domain features are extracted (B). These readings are projected onto a reduced feature space, where each authentication attempt is compared against previously learned thumprint expressions from group members (C). If a match, Thumprint will provide access by regulating an end-point such as an electronic lock (D).

Overview

Figure 11 shows a high-level description of how Thumprint works. To operate, Thumprint requires two components: a surface instrumented with an accelerometer and microphone (or a device already containing these sensors, such as a smartphone), and an endpoint to regulate access.

The *sensed surface* can take on many forms—e.g., a tablet touchscreen, a door, or tabletop. As a proof-of-concept implementation, we used an Android smartphone as our sensed surface. Meanwhile, the authentication *end-point* can be anything that regulates access control, such as a tablet or an electronic smart lock.

To use Thumprint, a group of at least two members must register themselves by entering the shared secret knock. To register, each member enters the secret knock on the sensed surface five to ten times. Thumprint records three-seconds of accelerometer and microphone data from each of the registration attempts, extracts a set of time- and frequency-domain features from those sensor streams, and stores each feature vector labeled with the individual's ID as training data. We selected a three-second duration to allow for sufficient variation in knock expression. Thumprint then processes these training data to “learn” both the shared secret knock and each individual's expression of the knock.

To later authenticate, an individual should reproduce the secret knock roughly in the same manner in which she registered. The system extracts an unlabeled feature vector from the authentication attempt and compares it against training data. If the unlabeled feature vector is similar enough to the group thumprint, it is authenticated as the member whose training data is most similar. Moreover, Thumprint computes a similarity score for each group member—so, depending on the security needs of the group, it is possible to provide tiered access control so that a knock is only authenticated if its similarity score is sufficiently high. If the score is too low, it is possible to provide lower tier access, or prompt the user to repeat the knock.

Training Pipeline

Once participants have provided a set of training data during the registration process, the key question is how can one use this training data to later authenticate group members? More formally, if we have an unlabeled authentication attempt, \vec{u} , we must determine whether or not to authenticate \vec{u} and, if so, which group member is most likely to have produced \vec{u} .

One approach is to use a one-class classifier, but these typically require a large amount of training data—dozens, if not hundreds of training points per group member. Instead, to make accurate decisions with fewer training data, we use a form of template matching: i.e., we compare \vec{u} to the set of templates, T , that are constructed during training to represent individual expressions of the shared secret knocks. If the distance between \vec{u} and any $\vec{t} \in T$ is sufficiently low, then we authenticate \vec{u} as coming from the user who produced \vec{t} . Otherwise, we reject \vec{u} as coming from an outsider. In brief, this process requires three implementation steps: *feature extraction*, *feature processing*, and *template construction*.

<i>Signal Transformation</i>	<i>Applicable sensor streams</i>	<i>Signal partitioning</i>	<i>Extracted features</i>
<i>Time-domain</i>	<i>Acceleration & Acoustic</i>	<i>Whole & One-second windows</i>	<i>Mean, mean absolute value, std. dev., max, min, RMS, zero-crossings, total energy, 2nd order average, third order average, average amplitude change.</i>
<i>Wavelets (D4)</i>	<i>Acceleration & Acoustic</i>	<i>Whole</i>	<i>Total power, max power, power bands, mean absolute coefficient value per band, coefficient standard deviation per band.</i>
<i>Fourier</i>	<i>Acceleration & Acoustic</i>	<i>One-second windows</i>	<i>Dominant frequency, spectral centroid, spectral rolloff, spectral crest factor, spectral flatness, lower 1kHz bins.</i>
<i>MFCCs</i>	<i>Acoustic</i>	<i>25ms windows</i>	<i>For each of the 12 coefficients, over all 25 ms windows: mean value, std. dev., mean first order-change, mean second-order change.</i>

Table 19. Features extracted for every thumprint, drawn from recommendations in prior work in sensing techniques. In total, 1020 features are extracted, though the feature space is dramatically reduced in later steps to avoid overfitting.

Feature Extraction

I extracted a set of features from each of the input acceleration and acoustic signals that users entered during registration. Features were extracted from the raw time-domain PCM values, as well as a Daubechies D4 wavelet and Fourier transformation (FFT) of the signals. For the raw-time domain and FFTs, I extracted features for each one-second segment of the signal to better preserve the temporal variance of the thumprints across the three-second window (i.e., to characterize thumprints that may be intentionally non-rhythmic and irregular). This was unnecessary for the wavelet transformation, as wavelet coefficients capture temporal variation by design [13]. Finally, for the acoustic signal, I also extracted features from the mel-frequency cepstral coefficients (MFCCs) computed for each 25 millisecond time-window of the signal. See Table 19 for an overview of features used.

At the end of the feature extraction process, we have a matrix, $F \in \mathbb{R}^{m \times n}$, where m is the number of training attempts in the system and n is the number of features that have been extracted. Each row of this matrix represents the features extracted for a particular training attempt. We also have a class vector, $\vec{y} \in \mathbb{Z}^m$, that represents which participant produced which row of F .

Feature Processing

Next, I employ a number of supervised pre-processing techniques on F . First, I use correlation-based feature subset selection (CFS) [50] to reduce the feature space to a parsimonious subset that distinguishes group members. The reduced feature space is reduced to *at most* one feature per row of training data to mitigate overfitting. I then discretize the feature space using Fayyad-Irani discretization [41]—a technique to bin continuous variables into discretized intervals that minimize the entropy of known class values in each bin. Supervised discretization can enhance predictive performance in many cases [41]. More intuitively, I discretize the feature space so that the template matching algorithm is less sensitive to micro-fluctuations in raw feature values. At the end of the feature selection and discretization process, we have a reduced matrix, $A \in \mathbb{Z}^{m \times k} = \text{Discretize}(CFS(F, \vec{y}))$, where $k \leq m$ is the number of features in the reduced feature space.

Template Construction

Following feature selection and discretization, we need to deconstruct the training matrix, A , into a set of known templates, T . The two most straightforward approaches are: (1) create a single template for each user by averaging all of their training attempts; and, (2) create a distinct template for each training attempt. However, both are suboptimal. The first approach fails to acknowledge that individuals might have multiple expressions of the shared secret knock—for example, one might sometimes enter the knock with more force, or other times at a slower pace. If all of these different

expressions are averaged, then the average will look different than any of the individual expressions. The second approach fails to learn common patterns across training attempts and reduces security by expanding the surface area in the feature space that represents the group shared secret. Thus, a single stray training attempt can compromise the security of the group by expanding the acceptable definition of the group's shared secret.

Instead, I take a middle-ground approach by clustering together related training attempts into distinct templates. With this compromised approach, we can detect multiple distinct expressions of the secret knock within users, but still minimize the surface area that represents the group shared secret in feature space. To do so, I run a k-means clustering algorithm on the training data for each individual group member and automatically determine the number of clusters that are appropriate using the average silhouette width method [91]. At the end of this process, we have a set of templates, T , that each contain a subset of the training attempts derived from one registered group member.

Authentication

Once training is complete, making an authentication decision on an unlabeled attempt, \vec{u} , is a matter of finding the cluster(s) closest to \vec{u} and then thresholding on the distance between \vec{u} and the closest cluster centroid:

$$\min_i d(\vec{u}, i) = \left(\sum_j \frac{|\vec{u} - \vec{t}_{ij}|}{k} \right) / |T_i|$$

where T_i represents the i th cluster and \vec{t}_{ij} represents the j th training vector in T_i , $|T_i|$ represents the size of T_i , and k represents the size of the feature space after feature reduction. In practice, the value of $d(\vec{u}, i)$ should typically fall within a range of 0 to 1 for any reasonably close attempt. Lower $d(\vec{u}, i)$ suggests a closer match between \vec{u} and T_i so in the simplest case of identification without authentication, we can identify \vec{u} as coming from the member who produced the cluster that minimizes $\min_i d(\vec{u}, i)$. To add authentication, we can introduce a threshold h . If $d(\vec{u}, i) \leq h$, then we authenticate; otherwise, we reject. Figure 12 visually illustrates the process.

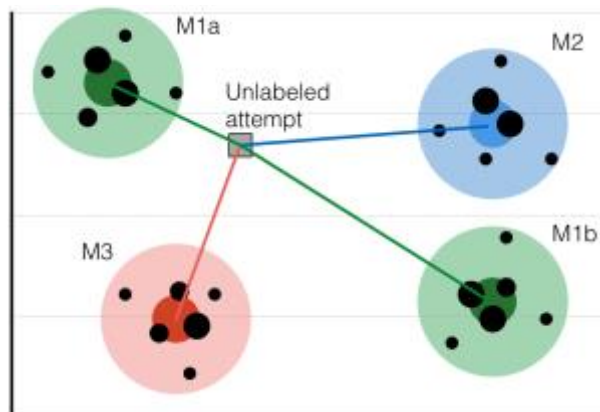


Figure 12. To authenticate, an unlabeled feature vector is transformed into the reduced feature space and then its distance to nearby training clusters is calculated. In this case, the unlabeled attempt would not be authenticated because it is too far from candidate clusters.

One potential concern is drift—or the idea that individuals might gradually change their expression of the secret knock over time. I handle drift by incrementally updating the training model as new training data is available (e.g., as a group member *successfully* authenticates over time) and by increasing the weight of more recent training attempts.

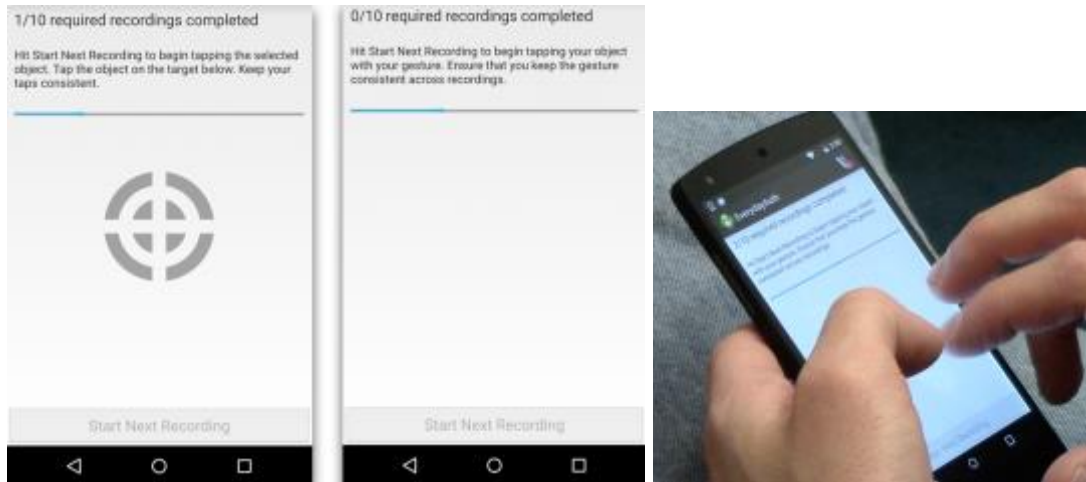


Figure 13. Screenshots of the app in which participants entered preset (left) and custom (middle) thumbprints. The right most figure shows how participants actually used the application interface.

Feasibility Evaluation

To evaluate the feasibility of my Thumbprint concept, I ran an initial lab study with 15 participants ranging in age from 18-55 years old (mean age 26, eight females). My goal with this feasibility evaluation was to answer the questions: Given a group of pre-registered users who all share a thumbprint and a set of un-registered adversaries who know the group's shared thumbprint, (i) how easily can outsiders impersonate group members? (ii) how often are group members confused as outsiders? And, (iii) how often are group members are confused for one another?

Procedure

To answer these questions, I ran a lab study. Consenting participants proceeded through two flows: a flow in which they entered pre-selected thumbprints, and a flow in which I had them create their own unique thumbprints. Participants entered their thumbprints on a Nexus 5 Android phone running custom software. For each thumbprint, my application recorded three-seconds of accelerometer data sampled at 2kHz and three-seconds of microphone data sampled at 44.1 kHz.

In the first flow, I selected 10 example objects that spanned a variety of materials: a *wooden* letter opener, a *rubber* eraser and fridge magnet, a *plastic* eye drop bottle, pen and chapstick, a *metal* Swiss army knife and watch, a *leather* wallet, and the participant's *knuckle*. Participants were instructed to hold the phone comfortably in their non-dominant hand. Then, for each of the 10 thumbprints, participants held the object in their dominant hand (or used the knuckle of their dominant hand) and knocked repeatedly on the center of the screen for three seconds. They repeated the entry of each thumbprint 10 times in total.

After completing this flow, participants were allowed to create their own custom thumbprints. Participants selected four tokens from the 10 objects provided and then had to develop their own unique knock for each of these tokens. Thus, participants could knock using any part of an object, anywhere on the screen and in any pattern. I demonstrated these options to participants prior to start of this flow. Participants again had to repeat each of their four unique thumbprints ten times each. I video-recorded participants entering their unique thumbprints so that I could later use these recordings to simulate shoulder surfing adversaries. Data from this flow was primarily used as raw material for the second study.

To improve data collection, the study interface provided a progress bar to inform participants of their three-second time limit. For the first flow, the interface also contained a target at the center of the screen to assist participants with their aim. Figure 13 shows screenshots of the process.

With 15 participants, 14 thumprints, and 10 repetitions per thumprint, the study yielded 2100 thumprints consisting of three-second accelerometer and acoustic streams. I computed the aforementioned time and frequency domain features for each of these instances.

Results

To answer questions (i), (ii), and (iii), I needed to simulate data from small groups with a shared thumprint, as well as outsiders attempting to break those thumprints. For the 10 pre-defined thumprints (first flow), simulating small groups and competent outsiders was straightforward. As each participant produced the same set of 10 thumprints, every participant could effectively be partnered with some number of other participants to simulate a small group, and every other participant could be a casual adversary.

Thus, I randomly aggregated different subsets of $n \in [3,5,10]$ participants to represent small groups of varying sizes. For training, I used a random sample of 80% of each group member's data, and kept a holdout set of 20% for testing. Then, for each simple thumprint, I used data from the remaining $15 - n$ participants to simulate a strong adversary who knew the group thumprint (as all users in the first flow entered the same thumprints).

It is worth noting that I did not design Thumprint to be extremely strong against adversaries who exactly knew the group thumprint. Yet, my results exceeded expectations.

Figure 14 shows the mean minimum feature vector difference, $\min_i d(\vec{u}, i)$, for authentication attempts by actual group members versus those of adversaries. From Figure 14, we can see a large and clear separation between the feature vector differences of authentic attempts ($d=0.32$) from adversarial attempts ($d=1.06$). In Figure 15, I plot the acceptance rate of these attempts as a function of a configurable authentication threshold. We can see that Thumprint worked well: at a threshold between $[0.5, 0.75]$, we achieved 100% true positives and no false positives.

This result is promising—suggesting that thumprints might provide reasonable outsider rejection while maintaining high insider acceptance. However, it is worth keeping in mind that the adversaries in this evaluation were not specifically trying to replicate a thumprint in a way that they observed someone else. Furthermore, I collected all data within a single session, so it is not surprising that people's testing attempts were quite similar to their training attempts. I address these weaknesses in

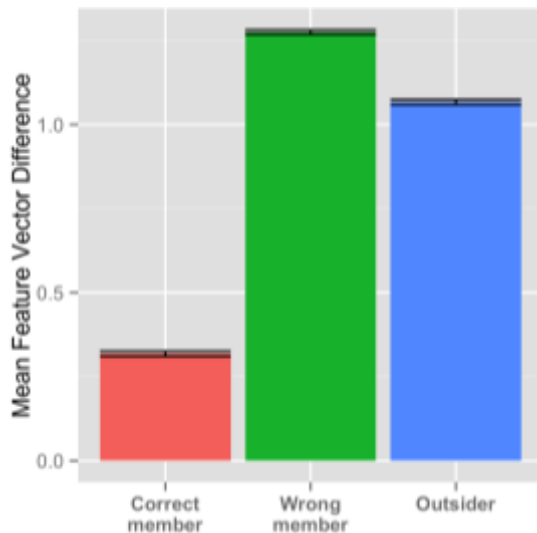


Figure 14. Mean feature vector difference (along with 95% confidence intervals) for user testing attempts (relative to their own training data and other group member training data), as well as outsider attempts.

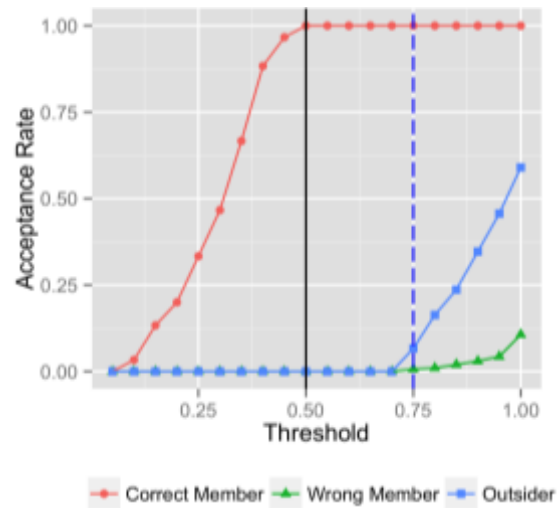


Figure 15. Acceptance rate as a function of feature vector difference. The black vertical line is where 100% of user attempts are accepted, and the blue dashed line is where >0% of outsider attempts are first accepted.

Part.	T1	T2	T3	T4	T5	T6
1	Main	S	V	V+T		V+T
2	Main	V	T	V+T		V+T
3	Main	T	S	V+T		V+T
4	Main	S	V	V+T		V+T
5	Main	V	T	V+T		V+T
6	S	Main	T	V+T	V+T	
7	V	Main	S	V+T	V+T	
8	T	Main	V	V+T	V+T	
9	S	Main	T	V+T	V+T	
10	V	Main	S	V+T	V+T	
11	V	T	Main		V+T	V+T
12	T	S	Main		V+T	V+T
13	S	V	Main		V+T	V+T
14	V	T	Main		V+T	V+T
15	T	S	Main		V+T	V+T

*Main: Group thumbprint; V: video+wrong token; S: sound only;
T: token only; V+T: video+correct token.*

Table 20. Study 2 flow for each participant. The columns represent the six thumbprints selected from Study 1. Cell values with “main” refer to thumbprints participants learned in session 1 and replicated in session 2. Other cell values refer to thumbprints replicated as adversaries.

my second study.

Consistency and Security Evaluation

I ran a second lab study, with 15 new participants, ranging in age from 18-57 years old (mean age 28, five females). My goal with this study was to answer the following two questions: (iv) can people consistently enter complex thumbprints after time-separated sessions? And (v) how well can thumbprint reject motivated adversaries?

Procedure

This study consisted of two 30-minute sessions that took place 24 hours apart. Broadly, I had participants register a thumbprint in the first session and re-enter the same thumbprint a day later. In addition, I had participants play the role of an adversary attempting to break into others’ thumbprints, given a set of capabilities and constraints.

Session 1: Participants initially had to enter four simple, pre-defined thumbprints to familiarize themselves with the application interface. This flow was the same as it was in the first study, where participants selected from a set of provided objects and tapped them repeatedly on the center of the screen. Once they had completed the pre-defined thumbprint flow, they were shown a video of a custom thumbprint created by a participant from the second flow in first study. Participants were allowed to watch the video as often as they liked. Once satisfied, they were instructed to replicate what they saw to the best of their ability. Participants were also told that they would have to re-enter this thumbprint the next day.

Of note, participants were shown one of three custom thumbprints corresponding to the study group to which they were assigned. I selected three groups because I wanted several participants to learn the same thumbprint so that I could later group them, and to ensure that the results were not tied to any single thumbprint.

Session 2: Participants came back for a follow-up session a day later. Their first task in this follow-up session was to re-enter the custom thumprint they had seen at the end of the previous day’s session. They had to do so from memory—no assistance was provided. Once completed, each participant had to enter four more custom thumprints. This time, however, I had participants play the role of adversary. Their task was to replicate other thumprints given a set of constraints to simulate different adversary models.

The four adversary models and their corresponding affordances were: (1) *video+correct token*: the full video recording of thumprint entry and use of the correct token; (2) *video+wrong token*: the full video recording of thumprint entry, but the correct token could not be used; (3) *sound only*: the audio recording of thumprint being entered (stripped from the video recording) and a best-guess attempt at picking the correct token; and, (4) *token only*: only knowledge of the correct token provided. Table 20 shows all of the thumprints each participant had to enter, along with the relevant constraints. Note that, as before, participants entered 10 repetitions for each thumprint.

At the end of the study, I had data for three thumprints (T1, T2, and T3) across two sessions from five participants each. For each of these thumprints, I also had 10 *video+wrong*, *token only* and *sound only* adversarial replications. For another set of three thumprints (T4, T5, T6), I had 10 *video+correct* adversarial replications. Notably, as *video+correct* adversaries can be considered authentic group members (if their data is included in the process of training Thumprint), I can divide the 10 *video+correct* replications into subsets of group members and adversaries as necessary.

Results

To answer the question (iv), can people remember and enter complex thumprints over time, I trained a model on data collected for T1, T2, and T3 from the first day’s session and tested it on data collected for those same thumprints collected in the second day’s session. Specifically, I calculated the minimum feature vector difference of the authentication attempts from the second session relative to data from the same user in the first session. As a point of reference, I also calculated the minimum feature vector difference of the 10 *video+wrong*, *sound only*, and *token only* adversarial attempts relative to group member training data from the first session. To see if group members

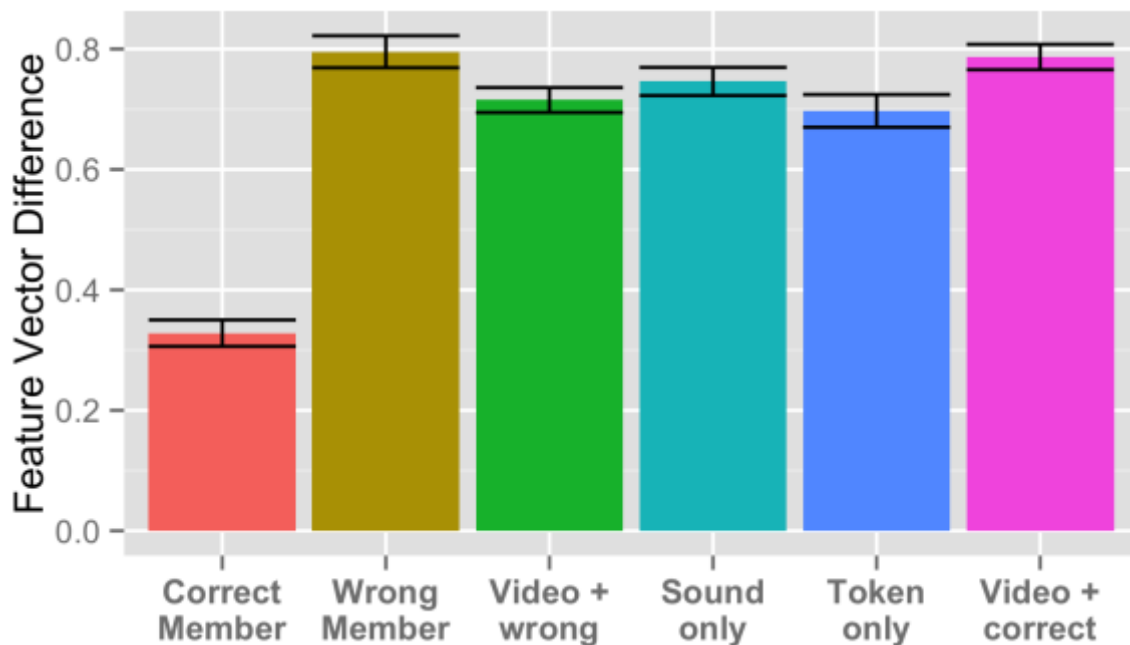


Figure 16. Mean feature vector difference (along with 95% confidence intervals) for T1-T3 across authentic and adversarial attempts. User testing data was collected one day after the training data.

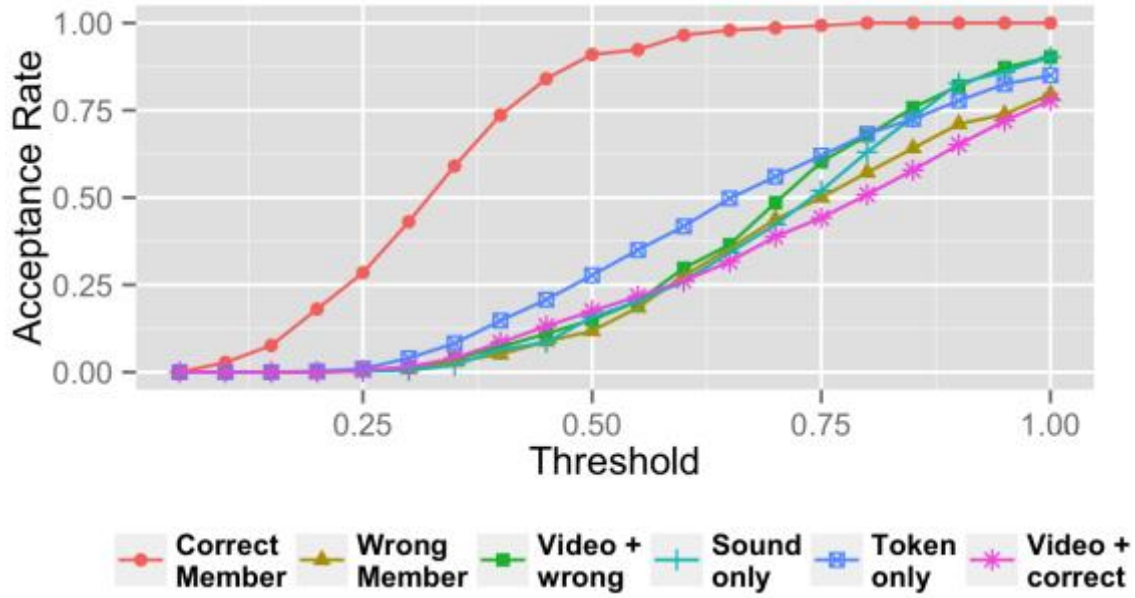


Figure 17. Acceptance rate as a function of minimum acceptable threshold across all thumprints. There is no threshold value to perfectly distinguish authentic attempts from adversarial attempts, but threshold values between 0.4 and 0.5 yield high true positives and low false positives.

could be misidentified with each other, I also calculated the minimum feature vector difference between user authentication attempts and the training data for other group members. Figure 16 shows the results.

We can see that mean feature vector difference for all authentication attempts by participants as compared to their own training data ($d=0.38$) from a previous session is much lower than the three adversary models ($ds=0.70, 0.74, 0.70$), as well as those of the wrong group members ($d=0.76$). In fact, participants are not much more inconsistent across time-separated sessions than they are within the same session ($d=0.32$ in Study 1). This marked difference between authentic user and adversarial attempts lends support to the conjecture that users can effectively replicate thumprints over time and cannot easily be impersonated by casual but motivated adversaries.

To definitively answer question (v), I next sought to translate these findings into individual authentication decisions. In addition to the models for T1-T3 that I used in the previous analysis, I also included models for T4-T6. Specifically, for each of T4, T5 and T6, I selected five participants to be “group members” and five participants to be *video+correct token* adversaries. I trained a model on 80% of the available data for the group members, holding out the additional 20% for testing.

Figure 17 shows a plot of acceptance rates for correctly identified group members (“correct member”), all four adversary types (*video+wrong*, *sound only*, *token only* and *video+correct*), as well as how often a user would be authenticated but misidentified as another member of the group (the “wrong member” trend line).

Expectedly, these results are not as optimistic as the analysis from my first study, when all data was collected from a single session and when the adversaries were not explicitly trying to exactly replicate the thumprint expression of a specific group member. One immediately notable result is that group members are rarely misidentified—this makes sense, as the preprocessing pipeline during training uses differences between group members to learn individual expressions of the thumprint.

However, adversaries can have some success at cracking thumprints, particularly at higher thresholds. A good compromise between false positives and false negatives appears to occur in between the threshold values of 0.45 and 0.5. In between those thresholds, authentic user attempts

are correctly let in between 85 and 91% of the time, while adversaries are granted acceptance between an average of 13% and 19% of the time. While these adversarial success rates seem high, they are comparable to other intentional behavioral approaches, such as TapSongs (83.2% user recognition, 19.4% adversarial acceptance) [112] and keystroke dynamics for user identification (83-92% recognition) [77].

Interestingly, what I believed was the “weakest” adversary model, the *token only* model, was most successful at cracking thumprints. This appears to be because adversaries with more information quickly honed in on how they would try to replicate the thumprint and simply repeated this process for all ten attempts. Token only adversaries, however, explored a wider space of possibilities with their 10 replications (i.e., they tried many different knocks as opposed to just one knock).

Finally, it is important to remember that Thumprint is not designed to provide perfect security against strong, motivated adversaries (who have advantages such as a video of the secret knock and ten unfettered attempts). I designed Thumprint to provide reasonable security, but emphasized inclusivity with identifiability, equitable distribution of responsibility and ease of sharing and revoking access. Indeed, for local group resources that are already largely physically secure (e.g., in homes), I believe these results suggest sufficient security.

It should also be noted that any probabilistic authenticator carries some risk of accidentally authenticating outsiders (e.g., even stronger, more sophisticated ones like Apple’s TouchID [85]). Indeed, given the similarity in outsider rejection performance between my approach, TapSongs [112] and RhythmLink [73], this detection rate could be a natural limitation of using sensible behavioral interactions for authentication—at least using existing sensors and modeling techniques. Still, I argue that this level of outsider rejection is reasonable for the small, local-group setting, especially given the focus on inclusiveness.

Discussion

The evaluations suggest that groups of users who enter the same thumprint can reliably be distinguished from one another; that users can enter their thumprints fairly consistently over time; and, that casual but motivated adversaries are often detectable and can thus be protected against. Taken together, these results suggest that Thumprint is a promising step towards the vision of socially-inclusive authentication for small, local groups. This evidence does not, however, suggest that Thumprint is immediately ready for mainstream use.

Though immediate viability is often an objective of traditional authentication research, I believe that this objective can be short-sighted. Traditional authentication works well for the purpose of identifying individuals who access private accounts, but Thumprint, and any other form of socially-inclusive authentication, is a significant departure from these models. Indeed, if the goal of traditional authentication is to create hard, impermeable boundaries that differentiate any two individuals, the goal of socially-inclusive authentication is to construct tweakable, semi-permeable boundaries between an in and out-group. While identifiability within the in-group is important, the process of identifying the individual should not raise hard barriers between those in the group.

Accordingly, while I have evaluated Thumprint to the standards expected of traditional authentication tools (e.g., with formally modeled adversaries), I believe this work opens up more interesting lines of inquiry. I reflect on some of these open questions and limitations, as well as discuss strategies for tackling them in future work.

Uncovering Hidden Group Authentication Needs

In designing Thumprint, I synthesized a number of unmet group authentication needs through a survey of the existing literature. However, these needs have only been explored in the socio-technical context of traditional authentication. As passwords and other typical forms of authentication have been long ingrained into everyday technology use, it may be difficult for users to conceptualize forms of authentication that are more group-friendly.

Accordingly, in future work, it would be pertinent to deploy Thumprint and other forms of socially-inclusive authentication as design probes in a field study with real groups. Through this field study, I

may uncover additional insights into how local groups use socially-inclusive authenticators and how they can be improved.

Designing for Group Variety

While Thumprint was designed to better cater to the authentication needs of local groups, these groups can have tremendous variety in their structure, composition and broader social context [109]. Families, for example, typically have little to no churn and often have clear power structures. Groups of friends, on the other hand, may be more egalitarian and prefer equal access to collectively shared resources. Work teams may have a lot of churn, be short lived, or require compatibility with broader security infrastructures. Student organizations may have expensive equipment that should be sharable, but require audit logs to keep track of who had access to what.

Many other factors no doubt affect how appropriate solutions like Thumprint are for groups. For example, some groups may have greater risk perception than others (e.g., a group of journalists). Other groups may be aversive towards probabilistic authentication as opposed to deterministic authentication. Still other groups may value anonymity and want to do away with identifiability, while preserving an equitable distribution of security responsibility.

Thumprint, thus, is likely to be better suited to the needs of some groups than others – it is not a panacea. Still, I believe it is a promising step forward and could be a starting point for further explorations into the design space of socially-inclusive authentication for different groups.

Strength of Security

Thumprint is not and was not designed to be perfectly secure. Though it is about as secure as comparable approaches for individuals (e.g., TapSongs [112], keystroke dynamics [77] and RhythmLink [73]), it is likely that a motivated adversary who observes individual group members entering the secret knock would be able to fool the model. Still, Thumprint's security may improve as more data from multiple time-separated sessions become available. As group members continue to use Thumprint for extended periods of time, there may be enough training data to employ these more sophisticated models (e.g., one-class classifiers) for stronger outsider rejection. In future work, I would like to explore this possibility.

Socially-Intelligent Interactive Security Systems

More generally, whereas the social proof nudges from Chapter 6 were an example of the first social prescription I prescribed, Thumprint is an example of the second: designing novel end-user facing security systems that are more social. Specifically, I designed Thumprint to quickly and easily authenticate and identify individual members of a small group with a single shared secret. Through two user studies, I found that individuals who enter the same thumprint can be reliably distinguished from one another, that people can enter thumprints consistently over time, and that Thumprint provides reasonable security against a variety of casual but motivated adversaries.

Thumprint is a promising first step towards the vision of socially-intelligent cybersecurity that better accommodates human social behavior in small group settings. It provides a degree of *inclusivity* that is atypical in traditional security systems and behaviors. Still, it is just a first step for a specific use-case. Thumprint is not the final form factor of social cybersecurity systems, but it is an illustrative example that social understanding and cybersecurity goals are not mutually exclusive.

Chapter 8: Discussion & Conclusions

Summary

The organizing question driving much of this work is: *How can we design systems that encourage better cybersecurity behaviors?* Engaging with this question is both important and urgent.

Designing systems that encourage better cybersecurity behaviors is important because security and privacy help realize the full potential of computing. Without authentication and encryption, for example, few would use digital wallets, social media or even e-mail. So it is unsurprising that the exploitation of weak security behaviors remains a massive enterprise. One estimate, calculated through a survey of over 3000 companies in the U.S., U.K. and Germany, suggests that the cybercrime industry is worth about \$450 billion dollars annually [49], with much of this value deriving from the exploitation of weak security behaviors—re-using passwords, ignoring software updates, neglecting two-factor authentication, etc. While the defenses necessary to combat much of today’s cybercrime already exists, few use those defenses in the way that experts recommend [60]. In other words, while existing defenses may be effective, their usage remains low.

Designing systems that encourage better cybersecurity behaviors is also urgent because cybercrime is emboldened by poor security habits. As Bruce Schneier argues, physical crime is largely prevented by social inertia [94]—that is, there are many social forces (e.g., family shame, legal reprimands, and stigma), that prevent burglary before a would-be burglar can ever get near one’s front door. Cybercrime has a relatively short history by comparison and the habits that make up the social inertia for cybersecurity are only being formed now. To date, these habits have been underlined by a sense of nonchalance: i.e., low awareness of security threats and available defenses, low motivation to act on security, and low knowledge of how to properly use security tools. This cannot continue. As computing encompasses more of our lives, we are tasked with making increasingly more security decisions. Simultaneously, the cost of every breach is swelling. Today, a security breach might compromise sensitive data about our finances and schedules as well as deeply personal data about our health, communications, and interests. Tomorrow, as we enter an era of pervasive smart things, that breach might compromise access to our homes, vehicles and bodies.

In this thesis, I have outlined one promising way in which we may be able to design systems that encourage better cybersecurity behaviors: by understanding and leveraging social influence. Social influence is known to be a big factor in human decision making, yet, prior to my work in this thesis, little was known about how it manifests in security decision making. To bridge that gap, I have done formative empirical work to construct an initial theory for social cybersecurity, combining both qualitative and large scale quantitative approaches.

In Chapter 3, I introduced an initial typology of social influences that affect security behaviors, finding that social influence plays a significant role in security decision making and that it can manifest in many ways. Chief among these is observability—when possible, people observe and emulate the security behaviors of others. Often, however, security behaviors are designed to be invisible or unobservable so it is difficult or impossible to spread through observability. In Chapter 4, I outlined the different types of conversations people tend to have about security, finding that people primarily speak to each other about security in order to warn or to teach, and that experts are often hesitant to share their knowledge for fear of being “boring” or sounding “preachy”. Finally, in Chapter 5, I presented an analysis of how having friends who use security tools affects one’s own likelihood to use those tools, uncovering the first large-scale empirical evidence that social influence affects security behaviors and that the design of a security tool strongly affects its potential for social spread. Specifically, security tools that are more *observable*, *inclusive* and *stewarded* are more amenable to social spread.

I then outlined two ways that this formative theory of social cybersecurity can be used to encourage better cybersecurity behaviors: (i) by constructing simple socially-inspired interface nudges; and, (ii) by implementing new end-user facing security systems that are more social by design. Exemplifying the first, I presented an experiment with 50,000 Facebook users showing that we can increase the adoption of existing security systems by increasing the observability of security tool usage among friends through simple notifications that provide social proof. Exemplifying the second, I presented

Thumprint: a novel, socially-inclusive form of authentication that can authenticate and identify members of a small, local group through a single shared secret knock. Taken together, my work suggests that social influence, when properly understood and leveraged in the context of cybersecurity, can indeed be used to encourage better cybersecurity behaviors.

This thesis provides both descriptive takeaways of how social influence affects security behaviors and prescriptive implications of how social influence can be used to improve end-user security. Next, I outline the most pertinent of these takeaways and implications.

Take-aways

Social influences strongly affect security behaviors

My thesis work provides some of the first empirical evidence that social influence strongly affects cybersecurity behaviors, and sometimes in unique and surprising ways that challenge expectations derived from social psychology.

Through my initial interview work (Chapter 3), I found that social influence accounted for nearly half of the recent security behavior changes made by my participants. Later, in the quantitative analysis of how security tools diffuse through social networks (Chapter 5), I found that social influence significantly affected the adoption of optional security tools on Facebook.

The mechanisms through which this social influence affects security behavior are manifold. Specifically, I distilled the following list of security-relevant social influences: **observing and emulating others' security behaviors**; **hearing about others' negative experiences**; **serendipitous teachable moments** (e.g., pranks and demonstrations); **collaborative sensemaking** (i.e., collectively discussing and making sense of security decisions and news events); and **sharing digital resources**. While likely non-exhaustive, this list offers an initial typology of how social influences affect security behaviors.

Notably, security system design, today, does not leverage these influences to improve security sensitivity. For example, while observability is one of the most intuitive and effective ways to spread good security behaviors, most security systems and behaviors are designed to be invisible. Thus there is a vast design space of social cybersecurity that remains largely untapped.

Security behaviors can have negative social consequences

As documented in Chapter 4 and in prior work by Gaw et al. [45], early adopters of security tools and behaviors are sometimes perceived as “paranoid”, “nutty” or as going “above and beyond” what is required. In other words, there is a social stigma to being overly cautious in the virtual world, just as there is in the physical world (e.g., the early perception of seat belts being uncool). This stigma has two strong negative effects.

The first negative effect is that the early adopters of security tools can create a disaffiliation effect where laypeople perceive good security behaviors as an indication that one is paranoid or has something to hide. In Chapter 5, I presented some empirical evidence that illustrates this negative effect of social influence: at low levels of exposure to friends who use standard security tools like Facebook's Login Notifications and Login Approvals, social influence has a negative effect on the further adoption of those tools. This negative effect did not, however, manifest for a more socially inclusive security tool: Trusted Contacts. Rather, the adoption of Trusted Contacts was positively affected by social influence even at low levels of exposure.

The second negative effect of this social stigma is that security experts often do not want to share their expertise. In the interview studies I presented in Chapters 3 and 4, for example, experts mentioned that they did not want to be perceived as “nagging” or “boring” so they did not often share their security knowledge. Unfortunately, experts presently have no outlet for sharing their concern for their friends and loved ones security that would portray them in a more positive light.

Existing end-user cybersecurity is often anti-social

Much of physical safety is rooted in sociality: for example, the ability to observe and emulate good safety behaviors (e.g., locking doors, wearing seatbelts) and the ability to implicitly offer protection to one and another through the idea of strength in numbers (e.g., walking home in pairs is safer than

walking home alone). Today's end-user facing cybersecurity systems fail to take advantage of sociality in at least three ways.

First, as I previously argued, is the lack of *observability*. Observability can help spread good security behaviors. In the physical world, we can often see threats to our physical safety and can observe how others respond to and protect themselves against those threats. In so doing, we learn a number of things: that there is a threat (raising awareness), that it is important to protect ourselves against that threat (raising motivation) and how to go about protecting ourselves from that threat (raising knowledge). In the virtual world, however, we are immersed in a fog of war. We cannot see the cybersecurity threats that may or may not be pertinent. We also cannot see how or even if others are responding to these threats.

Second is the lack of *inclusivity*. In the physical world, two people are generally physically safer together than they are apart because their strength aggregates: for example, it is generally easier for two people to fend off a burglar than just one person. In the virtual world, however, security does not aggregate in a group setting—instead, the strength of security for a group is only as strong as its “weakest” link. For example, if Alice has strong security behaviors and shares her files with Bob, who has weak security behaviors, the files Alice shares with Bob are only as secure as Bob's security. Thus, inclusivity in security is discouraged: experts are hurt by sharing data with laypeople, and laypeople do not benefit from being in a group with experts.

Third is a lack of *stewardship*. To use a more crude physical world analogue, it is sometimes easy to act on one's concerns for the safety of one's friends. For example, Bob can offer Alice a ride home at night if he believes it is not safe for Alice to walk home alone. It is easy for Alice to take Bob up on the offer, and it is relatively simple for Bob to provide Alice with a ride. In the virtual world, however, while there is a much larger divide between the knowledge and behaviors of experts versus non-experts than the physical safety differences between Alice and Bob, experts have no simple way to act on their concern for the security of their loved ones short of offering advice that can be perceived as nagging or boring.

Given this lack of observability, inclusivity and stewardship—all fundamental to our understanding and practice of physical security—perhaps it is no wonder that the general population's security sensitivity remains low.

The design of a security tool strongly affects its potential for social spread

In the analysis of how security tools diffuse through social networks I presented in Chapter 5, each of the three security tools I analyzed were affected by social influence differently. The adoption of standard security tools like Facebook's Login Notifications and Login Approvals was negatively affected by social influence at low levels of exposure to friends who use those tools. The effect eventually turned positive for Login Approvals at the highest level of exposure I tested, but remained negative for Login Notifications. For Trusted Contacts, however, the effect of social influence was positive throughout—i.e., even at the lowest level of exposure to friends who used Trusted Contacts, social influence had a positive effect on its adoption.

These different manifestations of the effect of social influence across different security tools suggest that the design of a security tool strongly affects its potential for social spread. Specifically, security tools that are more social by design are more likely to spread through social channels, whereas standard security tools that are asocial by design are more likely to only see use within early-adopter expert communities.

Greater exposure and diversity of exposure can counteract stigma

One commonality of the effect of social influence on the adoption of all three security tools I analyzed in Chapter 5 was that there was a positive main effect of exposure—that is, at higher levels of exposure to friends who use a security tool, the effect of social influence was increasingly positive. Another key finding from that study was the positive relationship between exposure to friends from diverse communities and the effect of social influence—i.e., people with exposure to friends from more distinct social contexts (e.g., high school, college, work) who use a security tool are more likely to use that tool than people with exposure to the same number of friends from fewer social contexts.

These effects were replicated in my social proof nudges experiment (Chapter 6), where I found that the effect of social proof was greater on people who had more friends who used a promoted security tool as well as on people who had friends from more diverse social contexts who used those tools.

Conversations about security are rare and are meant to warn or teach

In Chapter 4, I synthesized an initial list outlining the types of conversations people have about security and privacy. There were two key findings from that analysis. First, conversations about security and privacy are rare. Few people, not even experts, want to have face-to-face conversations about security. Laypeople are generally not interested except in specific circumstances, and experts do not want to seem “preachy” or “boring”, so will often only talk about security if the conversation is prompted by someone else. Then, what are these specific circumstances in which people have conversations about security? There are two: when someone wants to learn/teach, or when someone wants to warn. In other words, conversations about security are typically educational in nature and, so, only happen when a security behavior must be made or changed—e.g., when configuring a new device or in the wake of a security breach.

Security tools can be made more widespread by making them more social

In Chapters 3-5, I presented formative work on developing an initial theory for social cybersecurity. In Chapters 6 and 7, I presented two ways to use this theoretical foundation to improve end-user security. The first way is socially-inspired interface nudges. While it can be difficult to re-design existing security systems to be more social, it is possible to use simple interface nudges to make their usage more observable and inclusive. I presented an example of this in Chapter 6, by experimentally evaluating the effectiveness of notifications promoting the use of optional security tools with and without social proof on Facebook. In that experiment, the best social proof announcements attracted significantly more clicks than the non-social control, which, in turn, resulted in significantly more tool adoptions.

The second way is creating new end-user facing security tools that are more social by design, specifically emphasizing the design dimensions of observability, inclusivity and stewardship. In Chapter 7, I introduced Thumprint as an example of socially-inclusive authentication for small, local groups (e.g., families, small work teams). With Thumprint, I demonstrated that by relaxing the assumption that cybersecurity is meant to be an individual activity, it is possible to make end-user security systems that are social without compromising on security goals.

More generally, new security tools are held to incredibly high technical standards, but that is not always conducive to envisioning better futures. Immediate viability is a noble goal, but it should not be the only goal — doing so can be stifling and short-sighted and preclude more risky but fruitful alternative design considerations. Thumprint is an example — it is not perfectly secure, but it does provide an alternative socially-inclusive authentication design that is secure enough for many low-stakes use cases. Only by more fully exploring the design space of interactive security systems can we hope to find a design pattern that encourages better cybersecurity behaviors.

Open Problems and the Future of Social Cybersecurity

This thesis provides the first clear empirical evidence that social factors influence the uptake and spread of end-user facing security behaviors and systems. Through my work in modeling how social influences affect security behaviors and my work in applying those models to create more socially intelligent cybersecurity systems, I have laid the foundation for a new subfield that is related to yet distinct from usable security: social cybersecurity. This thesis, however, is just a launching pad. There remains a number of open problems and directions of inquiry that I expect will guide the future of social cybersecurity. Here, I discuss some of the more pertinent.

Discovering and validating the mechanisms underpinning social cybersecurity

One of the key findings of my thesis work is that social influence has a negative effect on the adoption of standard security behaviors and tools, like two-factor authentication. The *disaffiliation hypothesis* suggests that the reason for this negative effect is that the perceived early-adopters of security tools might be those who are perceived as “paranoid” or “expert” or otherwise unrelatably different to average end-users. In turn, as a result of this perceived unrelatability, the early adopters of standard

security tools might cast a stigma around the use of that tool. The disaffiliation hypothesis, however, is just a hypothesis. Accordingly, there remains a significant opportunity to explore and experimentally validate both the disaffiliation hypothesis and competing explanatory mechanisms for why there is a negative effect of social influence on the adoption of security tools.

Likewise, my work has only begun the exploration of applicable social psychology theory that can help explain end-user cybersecurity behavior. There remains much work to be done on uncovering the underlying mechanisms that drive security behaviors.

Cialdini, for example, lists seven principles of social influence: *reciprocity*, or that people tend to repay favors; *commitment* and *consistency*, or that people tend to honor commitments they explicitly make to others and to follow through with larger requests after agreeing to similar smaller ones; *social proof*, or that people tend to do things they see others do; *authority*, or that people tend to do as authority figures ask; *liking*, or that people tend to be persuaded by others they like; *scarcity*, or that people tend to value things that are believed to be scarce; and, *unity*, or that people tend to be more influenced by others with whom they identify [21,22]. Thus far, I have only experimentally validated the efficacy of social proof in the context of security behavior change. There remains a large opportunity to explore the (non)effectiveness of other principles of influence.

Another needed and important theoretical advancement is modeling how groups of people make joint security decisions. The pervasion of networked “smart” objects is increasingly making cybersecurity decisions salient in group settings. How do families decide on a shared password for a Nest thermostat account? How do employees in a shared working space decide on access control policies for communally owned paraphrenelia like mugs and cups? How do freelance work teams decide on whether or not to use secure messaging applications, and, if they do, which secure messaging application? Security has always been studied as an individual decision. But, as security will increasingly interfere with our social lives, it is becoming more important that we understand how security decisions are made in social contexts.

Finally, while I have shown that social influences can be used to affect behavior, it remains unclear whether and if there are differences in how those behaviors affect future security decisions. In other words, what, if any, are the differing effects, over time, between socially catalyzed behavior changes and non-socially catalyzed behavior changes.

Sociality as a third dimension in addition to usability and security

Instead of the traditional usability-security spectrum, my work on social cybersecurity suggests that we can instead think of usability, security and sociality as a three-dimensional design space for interactive cybersecurity systems. In other words, sociality is not usability. Indeed, something can be usable but not social (e.g., a graphical password), or social but not usable (e.g., an intelligent assistant that always makes the most socially appropriate authentication decision but takes 30 seconds to do so). Likewise, sociality is not security—we can have secure systems that are not socially intelligent (e.g., two-factor authentication), and socially intelligent systems that are not secure (e.g., a secret knock identification system that classifies a knock as definitely coming from one of a pre-registered set of group members). Rather, sociality is a third dimension that should be considered in the design of interactive cybersecurity systems.

Better understanding this design space, however, is a ripe opportunity for future exploration. Particularly pertinent is unpacking how sociality interacts with usability and security. Are there inherent trade-offs in sociality and security? Are there inherent synergistic properties between usability and sociality? As “usability”, “security” and now “sociality” are all complex concepts that cannot easily be formalized, these are questions that will need to be answered empirically through the construction and evaluation of many different “social” cybersecurity systems.

There is also a need to have a better understanding of what it means for a security system to be social and how its sociality can be expected to translate into security sensitivity. In this thesis, I have argued that there are at least three dimensions of sociality: observability, inclusivity and stewardship. An open problem is understanding how to make security systems that are observable, inclusive and stewarded. Another avenue for future work is to uncover and validate other social design

dimensions, beyond observability, inclusivity and stewardship. Ultimately, to inspire a new wave of socially intelligent cybersecurity systems, it would be pertinent to have a checklist or a set of guidelines that can help designers make security more social.

Still another avenue for future work is to better understand the boundaries of sociality in interactive cybersecurity system design. In other words, when is sociality important and necessary? When is it applicable? Usability, for example, is generally important but a lack of usability is more tolerable in high-risk situations (e.g., protecting national secrets). Usability, however, is of absolute importance in situations that are perceived to be lower-risk: e.g., in smartphone authentication, or in chat application encryption. Likewise, when is a lack of sociality more tolerable? When is it necessary?

Finally, more work is needed to understand how to make security for non-social contexts more social. It is relatively easy to provide users with notifications that their friends use two-factor authentication on Facebook than to, for example, provide users with notifications that their friends have public keys and prefer encrypted communications. The relative difficulty stems from the fact that there are few platforms, like Facebook, that have access to *both* security behavioral information *and* social connectivity information.

Metrics and methods for creating and evaluating social cybersecurity systems

More fundamentally, a large open problem is understanding the metrics and measures that can be used to evaluate social cybersecurity systems. If we make systems that are more observable, inclusive and stewarded, how can we translate those design goals into expected changes in security sensitivity. For example, what can we expect out of a more observable form of two-factor authentication? Perhaps the answer is X% higher adoption in a N-week timeframe relative to a non-observable form of two-factor authentication, or a Y% increase in awareness that two-factor authentication is an option. A goal for future work, then, would be to develop a better understanding of what those parameters X, N, and Y might be for different social designs. So far, however, we have relatively little understanding of how making security more social will translate into concrete, measurable behavior change.

What's needed is a set of concrete metrics and measures that can be used for evaluating the efficacy of a social cybersecurity system design, as well as standards for measurement. Just as usability has standard measures that are applicable to the design of usable security systems (e.g., the Nielsen Heuristics [79] or the System Usability Scale [15]), sociality needs a set of agreed upon measures and measurement instruments that are applicable to the design of social cybersecurity systems. Should we measure a sense of group cohesion or perceived social capital between group members? Higher group penetration of a security behavior among local groups of individuals? More discussion of security and privacy among family and other small group units? If so, what are reliable measurement instruments we can use to take these measurements.

These are just a few ripe opportunities and open problems for future work. My thesis has laid down a strong foundation for social cybersecurity, but there remains much to be done in order to develop a more holistic understanding of how social influences affect cybersecurity behaviors and how we can use those social influences to design systems that encourage better cybersecurity behaviors.

Conclusion

How can we design systems that encourage better cybersecurity behaviors? Despite years of improvements to the usability of interactive, end-user facing security systems alongside a rapid and sustained growth of cybercrime, many useful security systems remain underutilized. This trend cannot continue. As computing encompasses more of our lives, we are tasked with making increasingly more security and privacy decisions. Simultaneously, the cost of every breach is swelling. Today, a security breach might compromise sensitive data about our finances and schedules as well as deeply personal data about our health, communications, and interests. Tomorrow, as we enter the era of pervasive smart things, that breach might compromise access to our homes, vehicles and bodies. Accordingly, it is becoming increasingly important that security is something with which end-users actively utilize and engage.

One problem is that while the usability of security systems have improved, attitudes about the importance of end-user security have not—the awareness of security threats and available defenses, the motivation to utilize recommended security tools and behaviors, and the knowledge of how to use recommended security tools and behaviors remain low. This *security sensitivity* is unlikely to change through improvements to usability alone. Rather, attitude adjustments require longer-term social change.

Yet, to date, little theoretical work in usable privacy and security has applied social science theory to understand how social processes affect security sensitivity. In turn, this lack of theoretical insight has precluded systems work that accounts for the social consequences of security system design. Thus, there remains a great but largely untapped opportunity to model human social behaviors within the context of cybersecurity and in creating socially intelligent security systems that have a better understanding of these human social behaviors.

To bridge these gaps in theory and practice, in this thesis, I offered an initial theory of how social influences affect cybersecurity behaviors, distilled these theoretical insights into a set of broad design recommendations, and then implemented and evaluated two such systems that point to a promising future of social intelligent cybersecurity. My work provides key supporting evidence for the statement: **Social influences strongly affect cybersecurity behaviors, and it is possible to encourage better cybersecurity behaviors by designing security systems that are more social.** More generally, through my thesis work, I hope to have conveyed the following three points:

1. Social influence strongly affects cybersecurity behaviors, and the design of a security tool affects its potential for social spread. Specifically, security tools that are more *observable*, *inclusive* and *stewarded* are more likely to spread socially and spread beyond early-adopter expert communities.
2. It is possible to increase the awareness and adoption of existing security tools and behaviors by making their use more social through interface nudges—for example, notifications that offer people some social proof that others care about and act on on their own security.
3. There is a great but largely untapped opportunity to reshape interactive, end-user facing security systems to be more social—i.e., more observable, inclusive and/or stewarded—without compromising on key security goals.

References

1. Mark S. Ackerman. 2009. The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. *International Journal of Human Computer Interaction* 15, 2: 179–203. <http://doi.org/10.1207/S15327051HCI1523>
2. Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, et al. 2016. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *SSRN* 1, 40: 1–40. Retrieved from <https://ssrn.com/abstract=2859227/09-ART1%0Ahttps://ssrn.com/abstract=2859227/09-ART1>
3. Alessandro Acquisti and Jens Grossklags. 2004. Losses, Gains, and Hyperbolic Discounting: Privacy Attitudes and Privacy Behavior. In *The Economics of Information Security*, J. Camp and R. Lewis (eds.). 179–186.
4. Anne Adams and Martina Angela Sasse. 1999. Users are not the enemy. *Communications of the ACM (CACM)* 42, 12: 40–46. <http://doi.org/10.1145/322796.322806>
5. Devdatta Akhawe and Adrienne Porter Felt. 2013. Alice in warningland: a large-scale field study of browser security warning effectiveness. *Proc. USENIX Sec'13*, 257–272.
6. Sinan Aral, Lev Muchnik, and Arun Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *PNAS* 106, 51: 21544–9. <http://doi.org/10.1073/pnas.0908800106>
7. Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada Adamic. 2012. The role of social networks in information diffusion. *Proc. WWW '12*, ACM Press, 519–528. <http://doi.org/10.1145/2187836.2187907>
8. Albert Bandura, Joan E. Grusec, and Frances L. Menlove. 1967. Vicarious Extinction of Avoidance Behavior. *JPSA* 5, 1: 16–23. <http://doi.org/10.1037/h0024182>
9. Lujo Bauer, Lorrie LF Cranor, RW Robert W Reeder, Michael K MK Reiter, and Kami Vaniea. 2007. *Comparing access-control technologies: A study of keys and smartphones*. Retrieved from <http://repository.cmu.edu/cylab/46/>
10. Adam Beautement, M. Angela Sasse, and Mike Wonham. 2008. The Compliance Budget: Managing Security Behavior in Organisations. *Proceedings of the 2008 workshop on New security paradigms - NSPW '08*, ACM Press, 47. <http://doi.org/10.1145/1595676.1595684>
11. Mike Bond. 2005. The Dining Freemasons (Security Protocols for Secret Societies). In *Security Protocols*. Springer Berlin Heidelberg, 258–265.
12. Robert M Bond, Christopher J Fariss, Jason J Jones, et al. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489, 7415: 295–8. <http://doi.org/10.1038/nature11421>
13. Anders Brandt. 2011. *Noise and Vibration Analysis: Signal Analysis and Experimental Procedures*. John Wiley & Sons.
14. Cristian Bravo-Lillo, Lorrie F. Cranor, Julie Downs, et al. 2013. Your Attention Please: Designing security-decision UIs to make genuine risks harder to ignore. *Proc. SOUPS'13*.
15. John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability Evaluation in Industry* 189, 194: 4–7.
16. A J Bernheim Brush. 2012. It' s Used by Us: Family Friendly Access Control. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Workshop on Technology for Today's Family*.
17. Moira Burke, Cameron Marlow, and Thomas Lento. 2009. Feed me: Motivating Newcomer Contribution in Social Network Sites. *Proc. CHI '09*, ACM Press, 945–954. <http://doi.org/10.1145/1518701.1518847>

18. Damon Centola. 2010. The spread of behavior in an online social network experiment. *Science (New York, N.Y.)* 329, 5996: 1194–7. <http://doi.org/10.1126/science.1185231>
19. Damon Centola and Michael Macy. 2014. Complex Contagion and the Weakness of Long Ties. *American Journal of Sociology* 113, 3: 702–734.
20. Loren J. Chapman. 1967. Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior* 6, 1: 151–155. [http://doi.org/10.1016/S0022-5371\(67\)80066-5](http://doi.org/10.1016/S0022-5371(67)80066-5)
21. Robert B. Cialdini. 2009. *Influence*. Harper Collins.
22. Robert B. Cialdini. 2016. *Pre-Suasion: A Revolutionary Way to Influence and Persuade*. Simon & Schuster.
23. Robert B. Cialdini, Linda J. Demaine, Brad J. Sagarin, Daniel W. Barrett, Kelton Rhoads, and Patricia L. Winter. 2006. Managing social norms for persuasive impact. *Social Influence* 1, 1: 3–15. <http://doi.org/10.1080/15534510500181459>
24. Robert B Cialdini and Noah J Goldstein. 2004. Social influence: compliance and conformity. *Annual Rev. of Psych.* 55, 1974: 591–621. <http://doi.org/10.1146/annurev.psych.55.090902.142015>
25. Jacob Cohen. 1977. *Statistical Power Analysis for The Behavioral Sciences*. England: Lawrence Erlbaum Associates, Inc., Hillsdale, NJ.
26. L.F. Cranor and S. Garinkel. 2005. *Security and Usability: Designing Secure Systems that People Can Use*. O'Reilly Media.
27. Sauvik Das, Eiji Hayashi, and Jason Hong. 2013. Exploring Capturable Everyday Memory for Autobiographical Authentication. *Proc. UbiComp'13*.
28. Sauvik Das, Hyun Jin Kim, Laura A. Dabbish, and Jason I. Hong. 2014. The Effect of Social Influence on Security Sensitivity. *Proceedings of the 10th Symposium on Usable Privacy and Security (SOUPS'14)*.
29. Sauvik Das, Adam D.I. Kramer, Laura A. Dabbish, and Jason I. Hong. 2014. Increasing Security Sensitivity With Social Proof: A Large-Scale Experimental Confirmation. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*, ACM Press, 739–749. <http://doi.org/10.1145/2660267.2660271>
30. Sauvik Das, Adam D.I. Kramer, Laura A. Dabbish, and Jason I. Hong. 2015. The Role of Social Influence in Security Feature Adoption. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*, ACM Press, 1416–1426. <http://doi.org/10.1145/2675133.2675225>
31. Sauvik Das, Gierad Laput, Chris Harrison, and Jason I Hong. 2017. Thumprint: Socially-Inclusive Local Group Authentication Through Shared Secret Knocks. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, ACM Press, 3764–3774. <http://doi.org/10.1145/3025453.3025991>
32. Fred D Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and user Acceptance of Information Technology. *MIS Quarterly* 13, 3: 319–340.
33. Fred D Davis. 1989. Perceived Usefulness , Perceived Ease Of Use , And User Accep. *MIS Quarterly* 13, 3: 319–340.
34. Rachna Dhamija, J D Tygar, and Marti Hearst. 2006. Why phishing works. *Proc. CHI '06*, ACM Press, 581–590. <http://doi.org/10.1145/1124772.1124861>
35. Paul DiGioia and Paul Dourish. 2005. Social navigation as a model for usable security. *Proc. SOUPS '05*, ACM Press, 101–108. <http://doi.org/10.1145/1073001.1073011>
36. Paul Dourish, Rebecca E. Grinter, Jessica Delgado de la Flor, and Melissa Joseph. 2004.

- Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing* 8, 6: 391–401. <http://doi.org/10.1007/s00779-004-0308-5>
37. Serge Egelman, A.J. Bernheim Brush, and Kori M. Inkpen. 2008. Family accounts. *Proceedings of the ACM 2008 conference on Computer supported cooperative work (CSCW '08)*, ACM Press, 669. <http://doi.org/10.1145/1460563.1460666>
 38. Serge Egelman, Lorrie Faith Cranor, and Jason Hong. 2008. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. *Proc. CHI '08*, ACM Press, 1065. <http://doi.org/10.1145/1357054.1357219>
 39. Serge Egelman, David Molnar, Nicolas Christin, Alessandro Acquisti, Cormac Herley, and Shriram Krishnamurthi. 2010. Please Continue to Hold: An empirical study on user tolerance of security delays. *Proc. WEIS'10*. Retrieved April 25, 2012 from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.167.5560>
 40. Serge Egelman, Andreas Sotirakopoulos, Ildar Muslukhov, Konstantin Beznosov, and Cormac Herley. 2013. Does my password go up to eleven? *Proc. CHI '13*, ACM Press, 2379–2388. <http://doi.org/10.1145/2470654.2481329>
 41. Usama M. Fayyad and Keki B. Irani. 1993. Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning. *Proc. International Joint Conference on Uncertainty in AI*, 1022–1027. Retrieved from <http://trs-new.jpl.nasa.gov/dspace/handle/2014/35171>
 42. Adrienne Porter Felt, Alex Ainslie, Robert W Reeder, et al. 2015. Improving SSL Warnings. *Proc. CHI'15*, 2893–2902. <http://doi.org/10.1145/2702123.2702442>
 43. Adrienne Porter Felt, Robert W. Reeder, Hazim Almuhiemedi, and Sunny Consolvo. 2014. Experimenting at scale with google chrome's SSL warning. *Proc. CHI'14*, 2667–2670. <http://doi.org/10.1145/2556288.2557292>
 44. SM Furnell, A Jusoh, and D Katsabas. 2006. The challenges of understanding and using security: A survey of end-users. *Computers & Security* 25, 1: 27–35.
 45. Shirley Gaw, Edward W Felten, and Patricia Fernandez-Kelly. 2006. Secrecy, flagging, and paranoia. *Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI '06)*, ACM Press, 591–600. <http://doi.org/10.1145/1124772.1124862>
 46. Eric Gilbert. 2015. Open Book. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, ACM Press, 477–486. <http://doi.org/10.1145/2702123.2702295>
 47. Mayank Goel, Jacob Wobbrock, and Shwetak Patel. 2012. GripSense. *Proceedings of the 25th annual ACM symposium on User interface software and technology (UIST '12)*, ACM Press, 545–554. <http://doi.org/10.1145/2380116.2380184>
 48. Noah J. Goldstein, Robert B. Cialdini, and Vlas Griskevicius. 2008. A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. *Journal of Consumer Research* 35, 3: 472–482. <http://doi.org/10.1086/586910>
 49. Luke Graham. 2017. Cybercrime costs the global economy \$450 billion: CEO. *CNBC*. Retrieved from <http://www.cnbc.com/2017/02/07/cybercrime-costs-the-global-economy-450-billion-ceo.html>
 50. Mark A. Hall. 1999. Correlation-based Feature Selection for Machine Learning. *University of Waikato*. <http://doi.org/10.1080/01422419908228843>
 51. Alina Hang, Alexander De Luca, and Heinrich Hussman. 2015. I Know What You Did Last Week! Do You? Dynamic Security Questions for Fallback Authentication on Smartphones. *Proc. CHI'15*.
 52. James W. Hardin and Joseph M. Hilbe. 2007. *Generalized Linear Models and Extensions*. Stata

Press, College Station, Texas.

53. Eiji Hayashi, Sauvik Das, Shahriyar Amini, Jason Hong, and Ian Oakley. 2013. CASA: A Framework for Context-Aware Scalable Authentication. *Proceedings of the 9th Symposium on Usable Privacy and Security (SOUPS'13)*.
54. Eiji Hayashi and Jason I. Hong. 2011. A diary study of password usage in daily life. *Proc. CHI'11*, ACM Press, 2627. <http://doi.org/10.1145/1978942.1979326>
55. Cormac Herley. 2009. So long, and no thanks for the externalities. *Proc. NSPW '09*, ACM Press, 133–144. <http://doi.org/10.1145/1719030.1719050>
56. Cormac Herley. 2009. So long, and no thanks for the externalities. *Proc. NSPW '09*, ACM Press, 133. <http://doi.org/10.1145/1719030.1719050>
57. Cormac Herley and P van Oorschot. 2009. Passwords: If We're So Smart, Why Are We Still Using Them? *Proceedings of the 13th International Conference on Financial Cryptography and Data Security (FC'09)*. http://doi.org/10.1007/978-3-642-03549-4_14
58. Seong seob Hwang, Sungzoon Cho, and Sunghoon Park. 2009. Keystroke dynamics-based authentication for mobile devices. *Computers and Security* 28, 1–2: 85–93. <http://doi.org/10.1016/j.cose.2008.10.002>
59. Philip G. Inglesant and M Angela Sasse. 2010. The true cost of unusable password policies. *Proc. CHI'10*, ACM Press, 383–392. <http://doi.org/10.1145/1753326.1753384>
60. Iulia Ion, Rob Reeder, and Sunny Consolvo. 2015. “...no one can hack my mind”: Comparing Expert and Non-Expert Security Practices. *Symposium on Usable Privacy and Security (SOUPS)*, 327–346. <http://doi.org/10.1080/0888431022000070458>
61. Markus Jakobsson, Erik Stolterman, Susanne Wetzels, and Liu Yang. 2008. Love and authentication. *Proc. CHI '08*, ACM Press, 197–200. <http://doi.org/10.1145/1357054.1357087>
62. Maritza Johnson, Serge Egelman, and Steven M Bellovin. 2012. Facebook and Privacy: It's Complicated Categories and Subject Descriptors. *Proc. SOUPS'12*, 1–15. <http://doi.org/http://doi.acm.org/10.1145/2335356.2335369>
63. R. Collin Johnson. 2013. Cyber security solutions underused. *EE Times*. Retrieved from http://www.eetimes.com/author.asp?section_id=36&doc_id=1287251&page_number=1
64. Mike Just and David Aspinall. 2009. Personal choice and challenge questions: a security and usability assessment. *Proc. SOUPS'09*, ACM, 8. Retrieved November 17, 2011 from <http://dl.acm.org/citation.cfm?id=1572543>
65. Steven J. Karau and Kipling D. Williams. 1993. Social Loafing: A Meta-Analytic Review and Theoretical Integration. *Interpersonal Relations and Group Processes* 65, 4: 681–706. <http://doi.org/10.1037/0022-3514.65.4.681>
66. Amy K Karlson, A.J. Bernheim Brush, and Stuart Schechter. 2009. Can i borrow your phone? *Proceedings of the 27th international conference on Human factors in computing systems (CHI 09)*, ACM Press, 1647–1650. <http://doi.org/10.1145/1518701.1518953>
67. Brendan Kiley. 2005. Secret Knocks and Passwords. *The Stranger*. Retrieved January 5, 2017 from <http://www.thestranger.com/seattle/secret-knocks-and-passwords/Content?oid=25434>
68. Tiffany Hyun-jin Kim, Payas Gupta, Jun Han, et al. 2012. OTO: Online Trust Oracle for User-Centric Trust Establishment. *Proc. CCS '12*, ACM Press, 391–403. <http://doi.org/10.1145/2382196.2382239>
69. Ross Koppel, Sean Smith, Jim Blythe, and Vijay Kothari. 2015. Workarounds to Computer Access in Healthcare Organizations: You Want My Password or a Dead Patient? *Studies in*

- Health Technology and Informatics* 208: 215–220. <http://doi.org/10.3233/978-1-61499-488-6-215>
70. Adam D.I. Kramer. 2012. The spread of emotion via facebook. *Proc. CHI '12*, ACM Press, 767–770. <http://doi.org/10.1145/2207676.2207787>
 71. Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. 2010. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology* 10, 2: 1–31. <http://doi.org/10.1145/1754393.1754396>
 72. Bibb Latané, Kipling Williams, and Stephen Harkins. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology* 37, 6: 822–832. <http://doi.org/10.1037/0022-3514.37.6.822>
 73. Felix Xiaozhu Lin, Daniel Ashbrook, and Sean White. 2011. RhythmLink: Securely Pairing I/O-Constrained Devices by Tapping Felix. *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)*, ACM Press, 263–271. <http://doi.org/10.1145/2047196.2047231>
 74. Michelle L Mazurek, Brandon Salmon, Richard Shay, et al. 2010. Access control for home data sharing: Attitudes, needs, and practices. *Proceedings of the 28th international conference on Human factors in computing systems (CHI '10)*, ACM Press, 645–654. <http://doi.org/10.1145/1753326.1753421>
 75. M.B. Miles and M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. Sage Publications, Inc.
 76. Stanley Milgram, Leonard Bickman, and Lawrence Berkowitz. 1969. Note on the drawing power of crowds of different size. *JSPS* 13, 2: 79–82. <http://doi.org/10.1037/h0028070>
 77. Fabian Monroe and Aviel D. Rubin. 2000. Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems* 16, 4: 351–359. [http://doi.org/10.1016/S0167-739X\(99\)00059-X](http://doi.org/10.1016/S0167-739X(99)00059-X)
 78. Heidi Moore and Dan Roberts. 2013. AP Twitter hack causes panic on Wall Street and sends Dow plunging. *The Guardian*. Retrieved from <http://www.theguardian.com/business/2013/apr/23/ap-tweet-hack-wall-street-freefall>
 79. Jakob Nielsen. 1994. Heuristic Evaluation. *Usability Inspection Methods* 17, 1: 25–62.
 80. Deborah A. Prentice, Dale T. Miller, and Jenifer R. Lightdale. 1994. Asymmetries in attachments to groups and to their members: Distinguishing between common-identity and common-bond groups. *Personality and Social Psychology Bulletin (PSPB)* 20, 5: 484–493. Retrieved March 17, 2013 from http://books.google.com/books?hl=en&lr=&id=vI_k81fTpIAC&oi=fnd&pg=PA83&dq=Asymmetries+in+Attachments+to+Groups+and+to+Their+Members:+Distinguishing+Between+Common-Identity+and+Common-Bond+Groups&ots=l10cnkKvET&sig=X4r3XxpOGbweRedMyNHLtOBr0-o
 81. Emilee Rader, Rick Wash, and Brandon Brooks. 2012. Stories as informal lessons about security. *Proc. SOUPS '12*, ACM Press. <http://doi.org/10.1145/2335356.2335364>
 82. Elissa M. Redmiles, Amelia R. Malone, and Michelle L. Mazurek. 2016. I Think They're Trying to Tell Me Something: Advice Sources and Selection for Digital Security. *2016 IEEE Symposium on Security and Privacy (SP)*, IEEE, 272–288. <http://doi.org/10.1109/SP.2016.24>
 83. Elissa M Redmiles, Sean Kross, and Michelle L Mazurek. 2016. How I Learned to be Secure: A Census-Representative Survey of Security Advice Sources and Behavior. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16*, ACM Press, 666–677. <http://doi.org/10.1145/2976749.2978307>
 84. Karen Renaud. 2005. Evaluating Authentication Mechanisms. In *Security and Usability*, Lorrie

- Faith Cranor and S Garfinkel (eds.). O'Reilly Media, 103–128.
85. Frank Rieger. 2013. Chaos Computer Club breaks Apple TouchID. Retrieved January 5, 2017 from <https://www.ccc.de/en/updates/2013/ccc-breaks-apple-touchid>
 86. E.M. Rogers. 2003. *Diffusion of Innovations*. Free Press.
 87. E M Rogers. 1962. *Diffusion of innovations*. Free Press of Glencoe. Retrieved from <http://books.google.com/books?id=zw0-AAAAIAAJ>
 88. Everett M Rogers. 2003. *Diffusion of innovations*. New York, New York, USA.
 89. Daniel M Romero, Brendan Meeder, and Jon Kleinberg. 2011. Differences in the mechanics of information diffusion across topics. *Proc. WWW '11*, ACM Press, 695–704. <http://doi.org/10.1145/1963405.1963503>
 90. Paul R Rosenbaum and Donald B Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70, 1: 41–55.
 91. Peter J. Rousseeuw. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20, C: 53–65. [http://doi.org/10.1016/0377-0427\(87\)90125-7](http://doi.org/10.1016/0377-0427(87)90125-7)
 92. M.A. Sasse. 2003. Computer security: Anatomy of a Usability Disaster, and a Plan for Recovery. *Proc. CHI Workshop on HCI and Security Systems*, Citeseer. Retrieved February 21, 2012 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.9019&rep=rep1&type=pdf>
 93. Stuart Schechter, AJ Brush, and Serge Egelman. 2009. It's no Secret: Measuring the Security and Reliability of authentication via “secret” questions. *Proc. S&P'09*, Ieee, 375–390. Retrieved February 21, 2012 from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5207657
 94. Bruce Schneier. 2000. *Secret & Lies: Digital Security in a Networked World*. John Wiley & Sons.
 95. P Wesley Schultz, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vladas Griskevicius. 2007. The constructive, destructive, and reconstructive power of social norms. *Psychological science* 18, 5: 429–34. <http://doi.org/10.1111/j.1467-9280.2007.01917.x>
 96. Julian Seifert, Alexander De Luca, Bettina Conradi, and Heinrich Hussmann. 2010. TreasurePhone : Context-Sensitive User Data Protection on Mobile Phones. *Proc. Pervasive'10*, 130–137. http://doi.org/10.1007/978-3-642-12654-3_8
 97. Cosma Rohilla Shalizi and Andrew C Thomas. 2011. Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. *Sociological Methods and Research* 40, 2: 211–239. <http://doi.org/10.1177/0049124111404820>
 98. Steve Sheng, Bryant Magnien, Ponnurangam Kumaraguru, et al. 2007. Anti-Phishing Phil. *Proc. SOUPS '07*, ACM Press, 88–99. <http://doi.org/10.1145/1280680.1280692>
 99. Supriya Singh, Anuja Cabraal, Catherine Demosthenous, Gunela Astbrink, and Michele Furlong. 2007. Password sharing. *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '07)*, ACM Press, 895–904. <http://doi.org/10.1145/1240624.1240759>
 100. JM Stanton, P Mastrangelo, KR Stam, and Jeffrey Jolton. 2004. Behavioral Information Security: Two End User Survey Studies of Motivation and Security Practices. *AMCIS*, August: 2–8. Retrieved March 6, 2014 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.92.2938&rep=rep1&type=pdf>
 101. Eric Sun, Itamar Rosenn, Cameron A Marlow, and Thomas M Lento. 2009. Gesundheit! Modeling Contagion through Facebook News Feed. *Proc. ICWSM '09*.
 102. Xiaoyuan Suo, Y. Zhu, and G.S. Owen. 2005. Graphical passwords: A survey. *Proc. ACSAC'05*, IEEE. <http://doi.org/10.1109/CSAC.2005.27>

103. Richard H. Thaler and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth and Happiness*. Penguin Books.
104. Michael Toomim, Xianhang Zhang, James Fogarty, and James A Landay. 2008. Access control by testing for shared knowledge. *Proceeding of the Twenty-sixth annual CHI conference on Human factors in computing systems (CHI '08)*, ACM Press, 193–196. <http://doi.org/10.1145/1357054.1357086>
105. Amos Tversky and Daniel Kahneman. 1973. Availability : A Heuristic for Judging Frequency. *Cog. Psych.* 5, 2: 207–232.
106. Johan Ugander, Lars Backstrom, Cameron Marlow, and Jon Kleinberg. 2012. Structural diversity in social contagion. *PNAS* 109, 16: 5962–6. <http://doi.org/10.1073/pnas.1116502109>
107. Gérard Vincent. 1991. A history of secrets? In *A History of Private Life: Riddles of Identity in Modern Times*. 145–281.
108. Rick Wash. 2010. Folk models of home computer security. *Proc. SOUPS '10*, ACM Press, 1. <http://doi.org/10.1145/1837110.1837125>
109. Stanley Wasserman and Katherine Faust. 1994. *Social network analysis: Methods and applications*. Cambridge University Press.
110. Karl E. Weick, Kathleen M. Sutcliffe, and David Obstfeld. 2005. Organizing and the Process of Sensemaking. *Organization Science* 16, 4: 409–421. <http://doi.org/10.1287/orsc.1050.0133>
111. Alma Whitten and J.D. Tygar. 1999. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. *Proc. SSYM'99*, 14–28. Retrieved January 13, 2014 from http://www.usenix.org/events/sec99/full_papers/whitten/whitten.ps
112. Jacob Otto Wobbrock. 2009. TapSongs. *Proceedings of the 22nd annual ACM symposium on User interface software and technology (UIST '09)*, ACM Press, 93–96. <http://doi.org/10.1145/1622176.1622194>
113. Grace Wyler. 2013. AP Twitter Hacked, Claims Barack Obama Injured In White House Explosions. *Business Insider*. Retrieved from <http://www.businessinsider.com/ap-hacked-obama-injured-white-house-explosions-2013-4>
114. Yue Zhang, Serge Egelman, Lorrie Cranor, and Jason Hong. 2007. Phinding Phish : Evaluating Anti-Phishing Tools. *Proc. NDSS'07*.
115. 2014. Heartbleed. *Wikipedia*. Retrieved from <http://en.wikipedia.org/wiki/Heartbleed>
116. 2014. Vulnerabilities Summary for CVE-2014-1266. *National Vulnerabilities Database*.
117. 2016. Shibboleth. *Wikipedia*. Retrieved January 5, 2017 from <https://en.wikipedia.org/wiki/Shibboleth>