How I Learned to be Secure: a Census-Representative Survey of Security Advice Sources and Behavior

Elissa M. Redmiles, Sean Kross†, and Michelle L. Mazurek
University of Maryland †Johns Hopkins University

ABSTRACT

Few users have a single, authoritative, source from whom they can request digital-security advice. Rather, digital-security skills are often learned haphazardly, as users filter through an overwhelming quantity of security advice. By understanding the factors that contribute to users' advice sources, beliefs, and security behaviors, we can help to pare down the quantity and improve the quality of advice provided to users, streamlining the process of learning key behaviors. This paper rigorously investigates how users' security beliefs, knowledge, and demographics correlate with their sources of security advice, and how all these factors influence security behaviors. Using a carefully pretested, U.S.-census-representative survey of 526 users, we present an overview of the prevalence of respondents' advice sources, reasons for accepting and rejecting advice from those sources, and the impact of these sources and demographic factors on security behavior. We find evidence of a "digital divide" in security: the advice sources of users with higher skill levels and socioeconomic status differ from those with fewer resources. This digital security divide may add to the vulnerability of already disadvantaged users. Additionally, we confirm and extend results from prior small-sample studies about why users accept certain digital-security advice (e.g., because they trust the source rather than the content) and reject other advice (e.g., because it is inconvenient and because it contains too much marketing material). We conclude with recommendations for combating the digital divide and improving the efficacy of digital-security advice.

1. INTRODUCTION

It is easy to identify how most people learn the majority of basic skills: learning to ride a bicycle from parents, algebra from school, and how to make paper airplanes from childhood friends. The origins of users' security behaviors, however, are less obvious. The majority of today's users did not have parents or teachers who could instruct them about the dangers of computers and the internet. Instead, users collect digital-security advice haphazardly from a variety of sources including workplaces, the media, and stories of negative experiences that have happened to family and friends [9, 48, 49]. While a plethora of security advice is available from seemingly authoritative sources such as US CERT and Microsoft [1,2], and yet more is provided casually by friends and acquaintances, users adopt only a fraction of it. Further, it is unclear how useful and effective advice from these various sources may be, or whether the advice that is accepted is the most valuable.

Little effort, however, has gone toward understanding how and why users opt to adopt some recommended behaviors but reject others. Most prior research has instead focused on teaching users individual security-promoting behaviors such as phishing awareness [7, 11, 19, 22, 42, 51, 54, 57, 64]. A smaller set of prior work has hypothesized general models for understanding how security behaviors develop, but these models either have not been empirically validated [27] or have been based on small samples of 25 users or fewer [9,49].

In this work, we present the first large-scale empirical analysis of how users' security beliefs, knowledge, and demographics correlate with their sources of security advice, and how all these factors influence security behaviors. We find the first documented evidence of a digital divide with respect to security. In so doing, we investigate which users take advice from which sources, why they take that advice, and what they believe to be the purpose of security-relevant behaviors. Our results derive from a census-representative survey of 526 U.S. residents, which supports statistically generalizable conclusions regarding user learning, beliefs, and behavior [8]. Our survey queried respondents' behaviors, advice sources, reasoning, and beliefs across four digital-security domains identified as highly important by experts [33]: password strength, antivirus use, software updating, and two-factor authentication. To enable comparisons between digital security and the more well-developed domain of physical security, we asked similar questions about the securing of exterior doors in respondents' homes. We also collected participants' demographics, technical and security knowledge, and internet skill level. Our major findings include:

- **Digital Divide.** We are the first to document a concrete digital divide with respect to security. More specifically, we find that higher-skilled users, who tend to be socioeconomically advantaged, are significantly more likely to take advice from their workplace and to have the skills necessary to learn from a negative experience. In contrast, those with lower skills tended to take advice from family, friends and service providers

CCS'16, October 24 - 28, 2016, Vienna, Austria

ISBN 978-1-4503-4139-4/16/10...$15.00

DOI: http://dx.doi.org/10.1145/2976749.2978307

(e.g. TimeWarner). This divide may increase the vulnerability of already disadvantaged users, who in our study also self-report fewer security practices.

- **Advice Rejection.** Unsurprisingly, respondents most often rejected security behaviors because they were inconvenient. However, they rejected advice nearly as often for containing too much marketing material or because they had not yet had a negative experience. These findings support the need not just to make advice simpler and less intrusive, but also to minimize marketing messaging. They also confirm and extend prior work about the importance of negative experience stories [48], which we find to be not only an effective teaching tool but in some cases almost a prerequisite for certain security behaviors.

- **Advice Acceptance.** The majority of respondents accept advice based either on their evaluation of the advice content or their trust of the advice source. For passwords and physical security, respondents rely on their own evaluation of the content, perhaps because they feel comfortable with these concepts. With regard to updating, antivirus, and two-factor authentication, however, users are more likely to rely on their evaluation of the trustworthiness of the advice source. This may indicate a need to simplify advice on these topics in order to enable users to evaluate the advice content, independent of the source; alternatively, this may indicate the need to effectively signal source credibility.

Based on these results, we provide recommendations for combating the digital security divide and to carefully target a small set of critical advice, boosting its impact while reducing the burden of security advice overload.

2. RELATED WORK

In this section, we review prior work discussing factors that influence security behavior, the value of different security recommendations, and how users learn security behaviors.

2.1 Factors that Influence Security Behavior

Previous research has focused on demographic and social factors as well as beliefs and mental models that influence users' security behaviors. Howe et al. examined how the belief that one's information is not "important enough to hack" can influence security behavior, especially among low-socioeconomic status users [31]. Sheng et al. and Halevi et al. both examined the impact of personality factors on different security behaviors [23,56]. Das et al. found that social factors, such as peer-pressure, can make users more likely to adopt the same security behaviors as their friends [13, 14]. Similarly, negative experiences or stories from friends about negative experiences were found to affect users security decisions and beliefs [48]. Yet other work has shown the impact of mental models on users' security choices. Wash investigated how mental models of security influence users' perceptions of what is and is not dangerous [66]. Following on this initial qualitative work, Wash and Rader, in a large, census-balanced survey of U.S. internet users, found that those who were more educated tended to take fewer security precautions, despite having more advanced beliefs [67]. We extend this prior work to more generally evaluate the impact of demographic factors, technical background and exposure to security training, and security beliefs on how security behavior is learned.

A few generalized models have been proposed to explain users' security decisions. Herley suggests that users make a rational cost-benefit calculation when choosing what security advice or behaviors to accept [27]. Relatedly, Beautment et al. suggests a 'compliance budget' model, wherein users' limited resources that dictate their choice of security behaviors [9, 28]. However, the former model has not been empirically validated, and the latter was evaluated with a small, qualitative sample. We seek to extend this work by developing a generalizable quantitative model to explain user behavior as a factor of decision-making, advice sources, demographics, knowledge, and beliefs. We provide actionable results not only about why users practice different security behaviors, but how they learn these behaviors.

2.2 Evaluation of Security Recommendations

In any study of how users learn security behaviors, it is important to keep in mind which behaviors are most necessary for users' safety. In a recent survey, Ion et al. collected responses from more than 200 security experts regarding which security behaviors they most strongly recommend [33]. In our survey, we asked respondents about their behavior, decision-making, and beliefs with regard to four of the top suggestions from this paper: software updates, two-factor authentication, antivirus software, and using strong passwords. By studying how users learn security behaviors and find sources of advice, we can develop interventions to better enable the transference of these expert best-practices to general users.

2.3 Security Learning

In addition to investigating the impact of demographic factors and mental models on security behavior, and determining which behaviors are most important to learn, there has been some study of where users learn behaviors. Prior qualitative work concludes that users often learn from experience and the advice of family and friends [20, 21, 48, 49]. This paper builds on such qualitative work, especially our prior examination of users' advice sources, reasoning, and beliefs [49]. We build on those exploratory results, evaluating them at scale and generating statistically generalizable models of users' security-learning mechanisms that can be used in developing new educational interventions.

Additionally, work by Whitman examined user learning of security behavior in work environments [68], although little investigation has been done to verify whether workplace behaviors translate to the home environment. We expand upon this prior work with a larger and more representative sample and deeper inquiry into how users learn behaviors, not just from where they learned them. Further, we examine the transition of security knowledge from the workplace to the home computing environment.

A large body of work has focused on improving the efficacy of security behavior teaching tools. This work has touched upon improving phishing education [7,22,42,51,57], developing more effective warnings [6,16,55,62,70], reducing security-warning fatigue [10], and teaching users to create strong passwords [11, 19, 54, 64]. Specifically in the realm of advice, Rader and Wash recently examined security advice itself, using topic modeling to analyze connections between

user security decisions and the topics and words in three types of security advice [47]. However, to enable large-scale improvement in user security, we must go beyond examining the content of specific resources and improving these resources for individual behaviors. We need to develop a general model to understand how users learn digital security behaviors and consequently model the factors, such as advice sources and beliefs, that affect users' security. Such a model can enable the development and targeting of interventions to improve a user's security level in general, rather than behavior-by-behavior.

3. METHODOLOGY

We conducted a computer-administered, closed-answer survey of a census-balanced sample of 526 respondents in April 2016 via the Survey Sampling International panel. To ensure our survey instrument could produce generalizable and rigorous results, we pre-tested our questionnaire by conducting cognitive interviews and expert reviews. These methods are best practices in survey methodology for minimizing biases and improving validity in survey data collection [46].

This study was approved by the University of Maryland Institutional Review Board. Below, we discuss our survey development process, sampling procedure, details of our statistical analysis, and limitations of our work.

3.1 Survey Development

Our survey queries respondents': digital and physical security advice sources, reasoning for accepting or rejecting advice from these sources, beliefs about the purpose and value of different digital-security behaviors, and general opinions regarding the importance and utility of digital- and physical-security advice. In order to gain a sampling of respondents' advice sources and reasoning across different digital-security domains, we asked questions about their behaviors, advice sources, reasoning, and beliefs with regard to four behaviors highly recommended by security experts [33]: password strength, antivirus use, software updating, and two-factor authentication. To enable comparisons between digital and physical security, we asked similar questions about the securing of exterior doors in the respondent's home.

In addition to asking standardized demographic questions regarding respondents' age, race, gender, education level, and income, we also asked whether respondents had ever held a government security clearance, and if not, whether they currently work with data governed by HIPAA (U.S. health-privacy regulation) or FERPA (U.S. student-privacy regulation). We refer to these participants collectively as *sensitive-data participants*. Further, we asked whether respondents held a degree in or worked in the fields of computer science, computer engineering or IT. We also administered the six-item web-use skills index to assess the respondent's technical skill level [26]. We used this to explore how exposure to a security-sensitive mindset, educational background or work experience on computer science or IT, and technical skill, respectively, influence users' learning mechanisms and security behaviors.

3.1.1 Cognitive Interviews

After developing the initial set of questions, we conducted cognitive interviews with five demographically-diverse participants (see Table 1). Cognitive interviewing is a method of pre-testing questionnaires that provides insight into how respondents interpret and answer questions, so that errors can be corrected before deployment [46, 69].

Gender	Age	Race	Educ.	Income
M	40-49 yrs	Asian	B.S.	$100-$125k
M	18-29 yrs	Hispanic	M.S.	$30-$50k
M	30-39 yrs	Black	Some College	$30-$50k
F	50-59 yrs	Black	High School	<$30k
F	40-49 yrs	White	B.S.	$50-$75k

Table 1: Cognitive Interview Demographics.

During the interview, participants were instructed to "think aloud" as they answered each interview question via the Qualtrics survey interface. After answering each survey item, they were asked one of the following questions: "Was that question hard to answer?"; "How did answering that question make you feel?"; "Was there an answer choice missing or one that you would have preferred?" Participants frequently volunteered information about how they felt or missing answer choices, even when unprompted. The results of these interviews were used to iteratively revise and re-write certain questions until they were clearly understood by respondents. No participants reported finding the questionnaire stressful nor the questions uncomfortable [15]. The cognitive interviews were 20 to 30 minutes in length.

3.1.2 Expert Reviews

After the third cognitive interview was complete, three experts reviewed our survey instrument to evaluate question wording, ordering, and bias: our university's statistical and survey methodology consultant and two human-computer-interaction researchers with survey expertise. Expert reviewing is another best practice typically used to identify sensitive questions, questions that may need additional clarification, and problems with question ordering or potential biases [46]. We updated a number of our questions following the expert review, and then completed cognitive interviews until no additional questionnaire problems emerged.

3.1.3 Final Survey Instrument

The survey was administered via the Qualtrics web interface. Each question was required, and a "Prefer not to answer" choice was offered for any questions identified as sensitive by the researchers or the expert reviewers. Additionally, sensitive questions and questions that may have had social-desirability bias (in which the respondent may feel socio-cultural pressure to respond in an "acceptable" manner) were rewritten to reassure respondents that all answers were acceptable, according to best practices [63]. For example, a question asking whether or not respondents used antivirus software was phrased as follows: "There are different reasons that people decide to use or not to use antivirus software on their personal devices. Which of the following best describes you:...". A list of the final survey questions are provided as supplementary material.

The order in which the questions about each of the four digital security behaviors were presented was randomized to prevent order bias [41]. The physical-behavior question was not included in the randomization, as the results of the cognitive interviews showed that respondents needed to be prepared for the topic switch to physical security and found it cognitively challenging to switch between topics. In order to improve the quality of the data collected, a commonly-used attention check question was included: "Please select

Unhappy as your answer choice to the following question. This question is designed to check that you are paying attention" [59]. Finally, demographic questions were placed at the end of the questionnaire to minimize sensitivity and bias, as per expert recommendations and best practices [53].

3.2 Representative Sample

We wanted to examine a representative sample of internet-using adults in the United States. To this end, we followed the American Association for Public Opinion Research guidelines and used sample quotas to obtain obtain a census-balanced sample of US adults for our survey [8]. We contracted Survey Sampling International (SSI) to recruit respondents who matched the US-census sample quotas on the metrics of age, race, income and gender. SSI administered our Qualtrics survey to these respondents through their platform, and compensated respondents according to their agreement with SSI. Respondents were provided with benefits such as gift cards, airline frequent flyer miles, and donations to charities of their choice.

By using this large and representative sample, we can make statistically significant and broadly generalizable conclusions about user behavior, beliefs, and practices [8]. In comparison, the majority of work on user security behavior and learning is drawn from convenience samples on platforms such as Amazon Mechanical Turk and also from small qualitative lab studies [17, 33]. Prior work comparing Mechanical Turk samples with the general population has shown these samples to have important limitations when generalizing to the internet-using population both with regard to demographics and privacy attitudes [34,37,52]. As such, our work provides a more robust picture of user behavior at a national scale than has prior work.

3.3 Statistical Analysis

In addition to presenting descriptive statistics regarding the prevalence of respondents' behaviors, advice sources, reasoning, and beliefs, we compare Likert-scale factors between participant sub-groups (e.g. beliefs between sensitive-data and general participants) using the Mann-Whitney U test [43]. We also construct several logistic-regression models in an effort to describe the relationship between respondent security behavior and informal learning. Logistic regression is a well-known statistical method for modeling binary outcomes [30]. In order to avoid over-fitting these models, we used the standard backward-elimination technique, removing one factor from the model at a time, until we minimize the Akaike Information Criterion (AIC) [5,71]. We present the results of our models in Section 4.8. For each model, we present the outcome variable, included and eliminated factors, log-adjusted regression coefficients (*odds ratios*), 95% confidence intervals, and p-values.

We use Pearson's X^2 test to assess independence among categorical variables, such as between respondents' security behavior and their sources of computer security information [18]. For comparisons across many categories, we use omnibus tests; if the omnibus tests are significant, we then apply pairwise tests selected a priori to compare individual categories. One limitation of this method is that our data contains repeated measurements of the same respondent, as every respondent answered questions about multiple advice sources and security behaviors. Pearson's X^2 test does not take into account this repeated measurement, meaning it is possible reported test statistics are overstated; nonetheless we believe X^2 is the most appropriate test for these analyses. We interpret the results with this limitation in mind.

3.4 Limitations

As discussed in Section 3.2, our representative sample provides for robust, broadly generalizable results. It is currently not possible to obtain a purely probabilistic sample via an internet survey [8], as such we cannot precisely state the prevalence of user behaviors and advice sources in the entire U.S. population. Nonetheless, our work provides a strong foundation for understanding national behaviors and trends.

As with any survey, some respondents may have selected the first answer that seemed to satisfactorily answer the question, without thinking deeply about their own beliefs [40]. To mitigate this, we included an attention check question to screen out inattentive participants and kept our questionnaire to 10 minutes in length, following expert recommendations for minimizing respondent fatigue.

It is also possible that respondents mis-reported answers in an effort to answer in a socially desirable manner. However, we focus primarily on asking respondents to recall their advice sources, about which significant social desirability bias seems unlikely; additionally, our questionnaire-testing procedure revealed no evidence of social desirability bias. Further, while we asked participants to search their memory for answers to our questions, they may not have fully done so, or they may have forgotten some information. We also assume that participants are largely able to correctly identify which of their behaviors are security behaviors and why they practiced those behaviors. Finally, it is possible that our survey questions do not accurately assess the constructs we sought to measure. To mitigate these errors, we extensively pre-tested the questionnaire. While we made every effort to eliminate errors and biases, as with any survey, there may have still been lingering measurement errors.

4. RESULTS

In this section we describe our sample, the prevalence of behaviors in our sample, our findings regarding respondents' advice sources and beliefs, which users take advice from which sources, their reasons for accepting and rejecting advice, and how their sources of advice affect their security behaviors.

4.1 Sample

Our sample is nearly representative of the demographics of the United States with regard to age, education, gender, and race. Our sample is very slightly wealthier than the general population, potentially due to lack of internet or device access among those earning less than $30,000. Additionally, we had a 5% higher incidence of Caucasian respondents and a 5% lower incidence of Hispanic participants than in the general population. This may be due to the fact that we used a single "select all that apply" race and ethnicity question which offered both Hispanic (ethnicity) and White (race) as answer choices to the same question. Table 2 compares the demographics of our sample to the 2014 United States Census [3].

4.1.1 Knowledge and Skills

To assess the knowledge and skills of our respondents, we administered the extensively validated six-item web-use

Metric	Sample	Census
Male	49%	49%
Female	50%	51%
Caucasian	69%	64%
Hispanic	11%	16%
African American	12%	12%
Other	8%	8%
Some HS	3%	8%
Completed HS	23%	28%
Completed Some College	25%	18%
Associates Degree	10%	9%
College Degree	26%	26%
Master's	10%	7%
Doctoral	4%	4%
18-29 years	22%	23%
30-39 years	20%	17%
40-49 years	19%	17%
50-59 years	16%	18%
60-69 years	15%	14%
70+ years	8%	11%
<$30k	26%	32%
$30k-$50k	19%	19%
$50k-$75k	17%	18%
$75k-$100k	13%	11%
$100k-$150k	14%	12%
$150k+	9%	8%

Table 2: Demographics of participants in our sample. Some percentages may not add to 100% due to item non-response. Census statistics from the American Community Survey [3].

skills index, which measures internet skills on a scale from 1 (low) to 5 (high) [26]. The mean for our participants is 3.75 (SD=0.99). This is slightly higher than the mean of 3.37 that would be anticipated from Hargittai and Hsieh's work developing this scale. However, our result still seems reasonable, as Hargittai and Hsieh collected their data in 2010 and the internet skill level of the population has almost certainly increased in the past six years. Additionally, thirty percent of our respondents were 'sensitive-data' respondents and 19% of our respondents held a degree in or worked in the fields of computer science (CS), computer engineering, or IT.

4.2 Security Behavior

Overall, we find that 53% of respondents report making stronger passwords for some accounts than for others and 84% of respondents report using antivirus software. Although there has been no prior work requesting users to self-report their antivirus use, our findings agree with prior industry work by McAfee, which analyzed log data to determine that 88% of computers had antivirus software installed [58].

Updating. The majority of our respondents updated their software: 37% reported updating all of their software immediately and 41% reported updating their software a little while after learning of updates. Only 5% of respondents reported rarely or never updating their software, while 17% of respondents reported updating some but not all software. This self-report data contradicts the findings of Nappa et al., who used Symantec logs to find that at most 14% of vulnerabilities were repaired when an exploit was released [44]. Furthermore, our findings are also higher than those reported by Ion et al. who reported that 25% of experts and 9% of non-experts in their survey reported installing updates "immediately" [33]. These discrepancies could have several possible explanations: different sample (Nappa et al. measured machines rather than people; Ion et al. used Amazon Mechanical Turk while we used a census-representative sample), social desirability (although we believe our wording was more neutral), an increase in user updating behavior since the Ion et al. survey in 2014, differences in how different respondents interpret "immediately" updating their software, and/or the explicit "Updates are installed automatically" option in used in Ion et al's survey but not ours.

Two-Factor Authentication. Finally, of our respondents, 25% used 2FA on all of the devices or services that offered it; 45% used 2FA on some, but not all services; and 28% never used 2FA (2% NR). The proportion of our respondents that use 2FA is higher than the rates cited in prior work by Ion et al. [33]. This may reflect an increase in 2FA adoption since the Ion et al. survey was conducted in 2014; we also defined 2FA for all respondents, which may have prevented some under-reporting from participants who did not recognize the term. We asked the 236 (45% of total) respondents in our sample who used 2FA on some but not all services why they used 2FA for those services. The majority (62%) said they used 2FA on only some services because they were required to do so by those platforms. Twenty-eight percent of these 236 participants said that they used 2FA for the services that were more important to them. Very few participants (8%) said that they activated 2FA on only some services because it was easier to do so on those particular services. Of those who did not use 2FA for any services (149, 28% of total), 64% had never seen information about nor had been prompted to use this security strategy.

4.3 Beliefs about Behavior Purpose

To assess respondents' beliefs about the purpose of security behaviors, we asked respondents what they thought the "primary reason" was for updating software and for using 2FA. We did not ask these questions for passwords and antivirus use, as cognitive interview results showed that these behaviors had a single, intuitive answer, and asking a question deemed "obvious" by all respondents caused negative survey sentiment.

Security was not believed to be the primary driver for completing updates. The majority of respondents (40%) believed that the primary purpose was to "ensure the software is free of bugs and crashes less often." Twenty-nine percent of respondents selected "to increase the security of the software software" as the primary purpose, and 30% believed the purpose was to get the "latest and greatest software features" (1% NR). However, for 2FA, security was cited as the primary purpose by the majority, 67% of respondents. The remaining 21% of respondents believed that 2FA was used to ensure that they could regain access to their account, and 10% believed 2FA was to enable the website to contact them (2% NR).

4.4 Beliefs about Security Importance

In addition to asking respondents where they learned digital (and physical) security behaviors, we asked them to compare the usefulness and trustworthiness of digital and physical security advice using two 5-point Likert scales an-

chored on "Digital is a lot more useful (trustworthy) than physical" and "Physical is a lot more useful (trustworthy) than digital". Our prior qualitative work suggests that respondents who handle sensitive data value digital-security advice more highly [49]. In our sample, we found that sensitive-data respondents were significantly more likely to find digital-security advice more trustworthy than their non-sensitive peers (MWU $p<0.01$), while they were not quite significantly more likely to find digital-security advice more useful (MWU $p=0.08$) . We hypothesize receiving digital-security information and emphasis in the workplace leads to higher levels of trust of security information in general, and that implementing security practices in their routine workplace activities leads to sensitive-data respondents regarding digital-security advice as more useful.

4.5 How Users Learn Security Behaviors

For each behavior that a respondent reported completing, we asked how they learned the behavior or what instigated them to do the behavior. See Figure 1 for a summary of respondents' advice sources. Note that participants were allowed to select multiple options; as a result, percentages may add to more than 100%.

The majority of respondents (80%) cited device- or software-based prompts or requirements as a reason for doing at least one digital-security behavior. These prompts included password meters, update reminders, or invitations to use 2FA. Further, 53% of respondents cited being required to use a behavior or automatic behaviors, such as automatic software updates, as a reason for using a behavior.

Media and family/friends were the most prevalent sources of digital-security advice. This finding confirms the results of prior, less representative studies [21, 49]. For the majority of respondents, online, print, or TV news articles were at least one of the types of media advice they saw (67.5%). Online forums served as a digital-security advice source for 40% of respondents, and fictional narratives and advertisements accounted for 25% of media advice, each. Additionally, the majority (60%) of advice from a family members and friend was given by a person with a background in CS or IT. This confirms prior qualitative work, indicating that respondents may feel that these individuals are experts or individuals with these backgrounds may volunteer more unsolicited advice [45].

As also noted in prior qualitative work, negative experiences appear to be a key source of digital-security advice [48, 49]. Our work confirms this finding, although at a lower level of prevalence than may have been expected: 28% of our respondents learned to practice a behavior due to a negative experience or a story told about a negative experience.

We were surprised to find that digital-service providers, such as TimeWarner or Bank of America, were a source of advice for 33% of respondents, as this source of advice is little discussed in prior research. Respondents also reported receiving a significant amount of advice from their workplace (29.5%). Of those who received advice from work, over 50% received advice from an IT newsletter or a friend or colleague who worked in the IT department. The remainder received advice from formal security training or from colleagues who did not have an IT background.

4.6 Who Takes Which Advice?

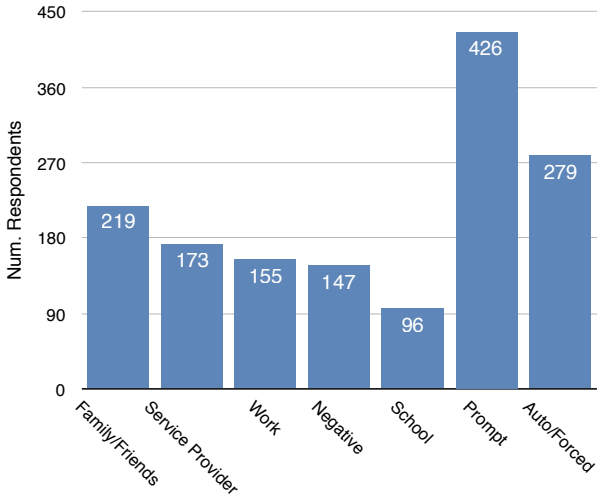

Figure 1: Prevalence of advice sources.

How do these beliefs about the usefulness and trustworthiness of digital-security advice, as well as demographic and knowledge factors, impact from where users take advice? In this section we present the results of binary logistic regression models for each source from which respondents reported learning behaviors. These models provide insight into the audience of each advice source. Recommendations for more appropriate targeting of advice and the improvement of online equity between users of different skill levels are discussed in Section 5.

Whether or not a respondent reported a given advice source at least once was used as the outcome variable in our models. The input factors considered in each model are listed in Table 3. Included in these factors are two interaction factors—between sensitive-data and beliefs about the usefulness and trustworthiness, respectively, of digital- vs. physical-security advice—which were informed by our prior qualitative work [49]. Below, we describe and interpret the significant factors included in each model. Because we did not conduct a controlled experiment, these results do not imply causality. The final regression results for each model after backward elimination, including nonsignificant factors, are shown in Table 4.

4.6.1 Media

As described in Section 3.3, we used backward elimination, minimizing AIC, to reach our final model. The final model for media advice included internet skill, exposure to sensitive data, age, income, and belief about the usefulness of digital security as factors. We find that respondents with higher internet skill are 32% more likely to use media as an advice source, potentially because media is increasingly being distributed online rather than through print, TV or radio.

4.6.2 Work

The final model for work advice included internet skill, exposure to sensitive data, age, income, education, and belief factors. Those who work with sensitive data are 4.5× more likely to cite their workplace as an advice source than those who are not exposed to sensitive data. This confirms results

Factor	Description	Baseline
Gender	Male, female or other.	Female
Age	18-39 years, 40-59 years, and over 60 years.	18-39 yrs
Income	<$50,000, $50,000-$100,000, and >$100,000)	<$50,000
Race	Black, Hispanic, White, and Other.	White
Education	Less than Bachelor's degree, Bachelor's degree, Graduate degree.	<B.S.
CS Background	Whether or not the respondent reported working in or holding a degree in CS or IT.	N/A
Sensitive Data	Whether or not the respondent reported working with HIPAA, FERPA, social security/credit card data, or holding an active or prior clearance.	No sensitive data
Internet Skill	Level of internet skill as measured by the six-item general web-use skills index [26].	N/A
Belief: Useful	Response to whether digital-security advice was **a lot** more useful than physical-security advice, somewhat more useful, equal, or that physical-security advice was somewhat or a lot more useful than digital. On a five-point Likert scale.	N/A
Belief: Trust	Response to whether digital-security advice was **a lot** more trustworthy than physical-security advice, somewhat more trustworthy, equal, or that physical-security advice was somewhat or a lot more trustworthy than digital. On a five-point Likert scale.	N/A
Sens. Data & Useful	Interaction between the Sensitive Data and Belief: Useful factors described above.	N/A
Sens. Data & Trust	Interaction between the Sensitive Data and Belief: Trust factors described above.	N/A

Table 3: Factors used in regression models. Categorical factors are represented by binary variable sets and individually compared to the baseline; a numerical value, centered on the middle value, was used for Likert factors.

of prior work: those who have clearances and/or handle sensitive data may receive benefits that improve their security through workplace training [49]. Those who had higher internet skill were 41% more likely to cite the workplace as a source of advice. Further, those who cited their workplace as a source of advice were more likely to believe that digital security was more useful than physical security. Thus, there is a need for increased digital equity and improved security interventions for users who do not have the opportunity to receive workplace training.

4.6.3 Negative Experiences

The final model to describe those who cited negative experiences, or stories of these experiences included internet skill and sensitive data. Those respondents with higher internet skill levels were 31% more likely to cite a negative experience as an advice source and those who handled sensitive data were 50% more likely. Those with higher internet skill or those who handle sensitive data may be more likely spend more time online, and thus may be more likely to have and learn from a personal negative experience. Furthermore, more skilled users, and those who are exposed to sensitive data, may be more likely to recognize a negative experience when it occurs and identify the underlying cause of that experience than less skilled or experienced users.

4.6.4 School

The final model for advice from school included age, CS background, income, sensitive data, and beliefs. Those with a CS background were approximately 6× more likely than those without this educational background to cite school as an advice source, and those who were over 60 were only 11% as likely as younger respondents to cite school as an advice source. This is likely due to the fact that computers were not used in schools until relatively recently, and those who work in the field of CS or IT and/or hold a degree in this field most likely received digital-security advice as they were obtaining their degrees or training.

4.6.5 Device Prompts

Perhaps because they have already learned digital-security behaviors from school, before they ever see a prompt, those with a background in CS were only 20% as likely as those without this background to cite a device prompt as a way that they learned about a particular security behavior. That said, those who used prompts as a source of security advice were more likely to believe that security advice was important—perhaps because this belief encouraged them to heed the security prompts. The final model for device prompts also controlled for internet skill and income.

4.6.6 Device automation/requirements

The model for device automation—that is, whether respondents reported learning about a security behavior because it was automated or required—included CS background, belief about the usefulness of digital-security advice, age, and education. Those with a background in CS or IT were only 59% as likely as those without this background to learn from security requirements or automations. Similarly to device prompts, this may reflect that respondents with technical education already know about security behaviors before encountering them as requirements. Additionally, those who used a behavior because it was automated or required were more likely to believe that digital security advice was important—a belief in the importance of security may inspire users to utilize prompts and automation.

4.6.7 Family and Friends

The final model for advice from family and friends included age, ethnicity, CS background, and the respondent's internet skill. There appears to be a relationship between age and taking advice from family and friends. Although there is no discernible pattern between the coefficients for ages 40-59 and over 60, there appears to be a significant difference between respondents who are 18-39 years old and those 40 and over. Additionally, respondents who are Hispanic were only 47% as likely to report taking advice from family and friends as White respondents. These findings indicate potential differences in how respondents in different age and ethnic groups chose to solicit advice.

4.6.8 Service Provider

Finally, the model for service provider advice included age, internet skill, exposure to sensitive data, CS background, gender, and beliefs about the usefulness of digital-security

Source	Factor	OR	CI	p-value
Media	Internet Skill	1.32	[1.09, 1.6]	< 0.01*
	40-59yrs	0.59	[0.39, 0.91]	0.02*
	Over 60yrs	0.87	[0.54, 1.41]	0.58
	$50k - $100k	0.85	[0.55, 1.31]	0.45
	> $100k	1.82	[1.12, 2.97]	0.02*
	Sensitive Data	1.50	[0.98, 2.29]	0.06
	Belief: Useful	1.09	[0.86, 1.38]	0.47
	Sens. Data & Useful	1.31	[0.94, 1.83]	0.11
Work	$50k - $100k	1.93	[1.16, 3.21]	0.01*
	> $100k	2.91	[1.63, 5.18]	< 0.01*
	Sensitive Data	4.53	[2.65, 7.74]	< 0.01*
	Internet Skill	1.41	[1.11, 1.8]	< 0.01*
	Sens. Data & Useful	1.58	[1.04, 2.41]	0.03*
	Belief: Useful	0.69	[0.48, 0.99]	0.04*
	40-59yrs	0.86	[0.53, 1.39]	0.54
	Over 60yrs	0.53	[0.29, 0.97]	0.04*
	Bachelors degree	1.53	[0.92, 2.53]	0.1
	Graduate degree	1.83	[0.97, 3.47]	0.06
Negative Experience	Internet Skill	1.31	[1.06, 1.63]	0.01*
	Sensitive Data	1.50	[1, 2.23]	0.05*
	40-59yrs	0.55	[0.35, 0.87]	0.01*
	Over 60yrs	0.61	[0.36, 1.04]	0.07
School	CS Background	6.22	[3.46, 11.17]	< 0.01*
	40-59yrs	0.26	[0.14, 0.47]	< 0.01*
	Over 60yrs	0.11	[0.04, 0.27]	< 0.01*
	$50k - $100k	0.57	[0.3, 1.09]	0.09
	> $100k	1.99	[1.06, 3.74]	0.03*
	Belief: Useful	1.14	[0.8, 1.64]	0.46
	Sensitive Data	0.99	[0.55, 1.78]	0.97
	Belief: Trust	1.23	[0.95, 1.6]	0.12
	Sens. Data & Useful	0.71	[0.46, 1.11]	0.13
Prompt	CS Background	0.20	[0.12, 0.35]	< 0.01*
	Belief: Useful	0.80	[0.65, 0.97]	0.02*
	$50k - $100k	0.48	[0.28, 0.83]	< 0.01*
	> $100k	0.64	[0.35, 1.19]	0.16
	Internet Skill	1.26	[0.99, 1.6]	0.07
Automatic	Belief: Useful	0.79	[0.67, 0.92]	< 0.01*
	CS Background	0.59	[0.37, 0.94]	0.03*
	Bachelors degree	0.60	[0.39, 0.92]	0.02*
	Graduate degree	0.68	[0.39, 1.18]	0.17
	40-59yrs	1.00	[0.66, 1.5]	1
	Over 60yrs	1.74	[1.08, 2.79]	0.02*
Family & Friends	40-59yrs	0.40	[0.26, 0.61]	< 0.01*
	Over 60yrs	0.51	[0.32, 0.83]	< 0.01*
	Black	0.66	[0.37, 1.21]	0.18
	Hispanic	0.47	[0.25, 0.89]	0.02*
	Other	1.56	[0.81, 2.99]	0.18
	CS Background	1.37	[0.85, 2.2]	0.2
	Internet Skill	0.87	[0.72, 1.05]	0.15
Service Provider	Sensitive Data	1.81	[1.21, 2.7]	< 0.01*
	Male	1.57	[1.07, 2.31]	0.02*
	Internet Skill	1.24	[1.01, 1.53]	0.04*
	40-59yrs	1.40	[0.9, 2.19]	0.14
	Over 60yrs	2.10	[1.27, 3.48]	< 0.01*
	Belief: Useful	1.14	[0.96, 1.35]	0.13
	CS Background	0.65	[0.39, 1.09]	0.11

Table 4: Regression results for advice source models. OR is the odds ratio between the given factor and the baseline; CI is the 95% confidence interval; statistically significant factors (p<0.05) are denoted with *.

advice. Respondents who are male were 57% more likely to report receiving advice from service providers, than Female respondents. Additionally, respondents with higher internet skill levels were 24% more likely to report taking advice from a service provider, and those who are exposed to sensitive data were 81% more likely to have taken advice from this source.

4.7 Why Users Accept and Reject Advice

In addition to examining the factors that affect which users take advice from which sources, we wanted to understand why the users of each advice source choose to accept and reject advice and behaviors. Prior work has identified a number of different reasons that users accept and reject security advice [49]. In order to evaluate these findings and

determine their prevalence, we asked respondents why they chose to practice (or not practice) a behavior based on the advice that they received. The answer choices that we provided were drawn from our prior qualitative work and feedback gained during our cognitive interview sessions [49]. We detail the results below.

4.7.1 Advice Acceptance

To assess respondents' reasons for accepting advice, we asked, "Which of the following best describes why the information you received made you decide to do this behavior?" We then presented four answer choices. The first two "I trusted the person or source of the information." and "The information made sense to me." were drawn from our prior work [49]. The other two choices, "The information increased my fear of a negative event." and "Other." (with a write in option), were added based on the cognitive interviews and expert reviews to ensure that we captured all possible respondent answers.

From our prior work, we hypothesized that respondents would be more likely to accept physical-security advice based on their evaluation of the advice content, while they would accept digital-security advice based on their trust of the source. As shown in Figure 2, we found that trusting the source was most popular for antivirus (53%) and updating (51%), while trusting the content was most popular for 2FA (52%), passwords (57%), and door locking (58%). This appears to confirm our hypothesis for some digital-security behaviors, but not others.

To investigate further, we ran an omnibus X^2 across all advice sources and behaviors ($X^2 = 58.96$, $p = 4.79e{-}12$), followed by planned pairwise comparisons of each digital behavior to door locking. [1] Our results strongly support that most digital behaviors are different from the physical behavior, especially antivirus ($X^2 = 41.15$, $p = 1.41e{-}10$) and updating ($X^2 = 25.49$, $p = 4.45e{-}7$). Two-factor was also significantly different from physical, but to a lesser degree ($X^2 = 6.78$, $p = 0.0083$). We hypothesize that passwords are not significantly different from physical ($p > 0.05$) because participants have been exposed to enough passwords advice to feel comfortable evaluating its content directly.

Finally, we validate that the two choices presented in our prior qualitative work [49] are near-exhaustive of the reasons that users accept advice: for each of the behavior questions, only a small portion of respondents (μ= 5%) reported that increased fear of a negative event caused them to take advice; an average of 2% of our respondents selected "Other."

4.7.2 Advice Rejection

Similar to advice acceptance, we drew the answer choices for our question regarding why respondents chose *not* to practice a behavior, even after seeing information recommending that behavior, from multiple prior studies [48, 49, 66, 67] and from the results of our survey pre-testing. We only asked these questions regarding antivirus, updating, and 2FA; as it is rarely, if ever, an option to not use passwords. Nearly half of our respondents (43%, 225) rejected at least one of these three behaviors. Among these respondents, we found that inconvenience (28%) and advice that

[1] As described in Section 3.3, the significance of the results in this section are somewhat overstated due to repeated measures; however, the results are so strong that we believe they hold.

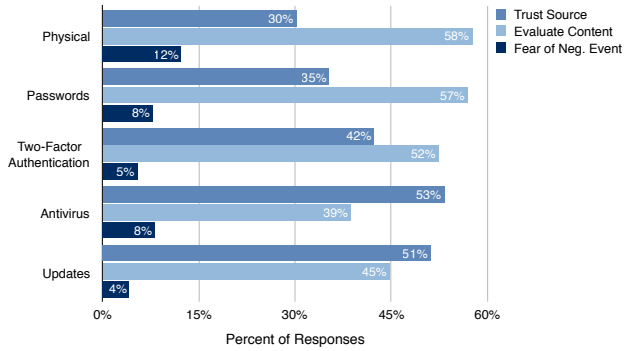

Figure 2: Reasons for accepting digital-security advice. Percentage per behavior.

contained too much marketing material (17%) were the two most common reasons for advice rejection, across all behaviors. We also found that a *lack* of negative experience was the third most common reason (13%) for rejecting a behavior. Although believing that one's data has no value [31], difficulty understanding advice [4], and being 'careful' on the internet [66, 67] have been offered as reasons for rejection in prior work, these reasons were all cited by less than 10% of our respondents.

The reasons respondents selected for rejecting advice varied by behavior.

Antivirus. For antivirus software, "They were trying to sell me something" was most often cited as a reason for rejecting advice related to antivirus software (33%). Other reasons for rejection included having had no prior negative experience (13%), feeling that they were "careful" when using their computer and the internet (15%), and finding antivirus software too difficult to use (11%).

Updating. Inconvenience, which included both "it was inconvenient" and "I did not have time", was the most common reason (50%) selected by respondents for why they did not complete software updates. All other reasons provided were cited by fewer than 10% of respondents and are thus not reported here.

Two-Factor Authentication. Inconvenience was also the most common reason given by respondents for not using 2FA (41%). Our prior work also suggested that privacy concerns may inhibit advice taking, especially for 2FA, where users may be reluctant to share a phone number. We found that while privacy concerns were not very prevalent in general, they were somewhat more prevalent for 2FA (15%) [49]. See Figure 3 for more detail on respondents' reasoning for rejecting security advice.

4.8 Advice Sources and Behavior

We next examine which advice sources were most commonly associated with which security behaviors (Figure 4). We find that media was the primary source of advice for both passwords (28%) and 2FA (26%), while family and friends accounted for a larger portion of the advice about antivirus behavior (28%) and software updates (24%). Service providers (21%) were also a primary advice source for 2FA. Finally, learning a behavior via a negative experience was most common for antivirus use (19%). An omnibus X^2 test showed that these differences among behaviors are significant ($X^2 = 59.05$, $p = 3.67e{-}7$).[1]

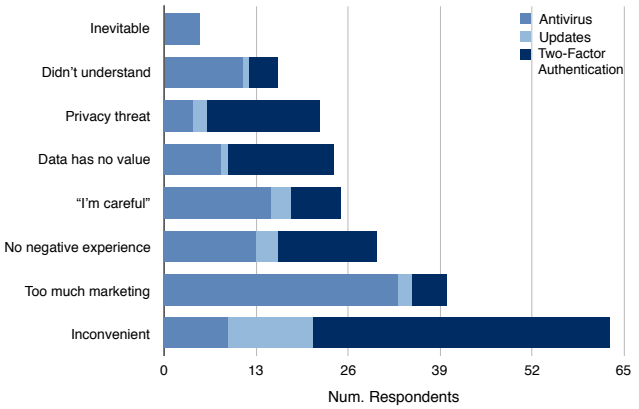

Figure 3: Reasons for rejecting digital-security advice. Total per behavior, multiple responses possible. This question was not asked for passwords, as not using them is rarely, if ever, an option.

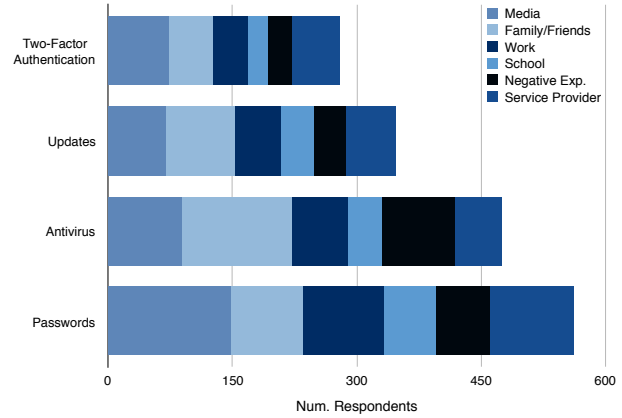

Figure 4: Advice source prevalence by behavior.

5. DISCUSSION

Below we draw conclusions from our findings and suggest directions for future research and design.

Digital Inequity. Users with lower socioeconomic status tend to be part of a *knowledge gap*: they have diminished access to digital media and more difficulty finding reputable and useful information on the web [24, 25, 50, 61, 65]. Our work expands these findings to digital security: we find evidence that users with higher levels of internet skills—demonstrated by prior work to be wealthier and somewhat more secure [26, 38]—use different advice sources. In particular, lower-skill users rely more on prompts, the advice of family and friends, and service providers than higher-skill users do. These differences, combined with discrepancies in skills and resources, may lead already disadvantaged users to be disproportionately victimized [12, 35]. Indeed, while we could not control for all confounding factors, we found that users with lower incomes were less likely to update their software (X^2=28.03, $p \leq 0.001$), to use 2FA (X^2=15.60, $p = 0.004$), and slightly less likely to use stronger passwords for sensitive accounts (X^2=9.60, $p = 0.048$). Although prior work has suggested that differences in security behaviors may be caused by lower-SES users not highly valuing their data [31], in our sample this was not the case.

Our respondents were 41% more likely to cite their workplace as an advice source if they had higher internet skill, and 4.5× more likely if they held a job that we categorized as security-sensitive. While the workplace may be a valuable source of advice for those who have access to it and the skills to understand this training, such resources may not be available for low-SES users; furthermore, security is often forgotten in digital literacy interventions that do exist in the workplace. For example, the Kesla+ project aimed at increasing the digital skills of low-skill office workers in the workplace included no training on digital security [39]. Thus, we advocate piloting and evaluating digital literacy programs which include or focus entirely on digital security. Additionally, future work should include analyzing and improving the grade-level readability and clarity of security advice to avoid widening the digital security gap.

We also found that participants with higher levels of internet skill were more likely to have learned from a negative experience, either their own or someone else's. We hypothesize that lower-skill users are less likely to recognize the causes of a negative experience and therefore learn from it. Because our results indicate that users who have not learned from negative experiences are more likely to reject advice, this inequity may put lower-skill users at additional risk. Of course, we want to minimize all users' direct exposure to negative experiences; instead we recommend amplifying stories of others' negative experiences. Future work could examine how to effectively simulate negative experiences, for example by using short, relatable stories that clearly demonstrate how to prevent the problem.

Advertisements & TV Shows. Our findings suggest the need for additional research into the content and teaching power of advertisements and TV shows. Fifty-percent of respondents who reported receiving digital-security advice from media received that advice from ads and TV shows. Thus far, the research community has had little input into nor focus on these forms of media. We recommend leveraging work in the communications and health fields [29, 32, 36, 60] to rigorously evaluate this media.

Improving Security Advice. Finally, we make two specific suggestions for improving digital-security advice. First, we find that much advice from the workplace and from family and friends comes from those with backgrounds in IT and CS. Anecdotally, many people with technical backgrounds are overloaded with "help-needed" requests from friends, family, and colleagues. Thus, we suggest that schools and the companies that employ these individuals support this important but potentially under-rewarded role, by providing more resources to help them meet these requests effectively. Disseminating a small set of essential security advice via these channels could have a large positive impact on user behavior. Second, we recommend explicitly stating the source and purpose of security advice. Nearly 50% of users accept advice because they trust the advice source; thus, it is crucial that the source be clearly identified and demonstrably authoritative, for example via professional credentials.

6. SUMMARY

In this paper we presented the results of a survey of 526 census-representative U.S. users' beliefs, behaviors, and sources of advice for digital and physical security. We find the first explicit evidence of a digital security divide: users with

higher skill levels and socioeconomic status, as well as those who handle sensitive data at work, are significantly more likely to get advice from work and to learn from negative experiences. This divide may increase the vulnerability of already disadvantaged users. In addition, we confirm that users evaluate advice based on the credibility of the source (rather than the content of the advice) significantly more often for digital-security topics than for a physical-security topic. Our results indicate that users reject advice not only because it is inconvenient and they have maxed-out their compliance budget, but because it contains too much marketing material; recommendations to use 2FA are also notably rejected due to privacy concern. Based on these results, we provide recommendations for combating the digital security divide, as well as recommendations for increasing the impact of security advice more generally. We hope that these results can improve the efficacy of truly critical recommendations, by helping them to stand out within the noisy space of other advice.

7. ACKNOWLEDGMENTS

Our thanks to Guy Aurelien, Lujo Bauer, Brenna McNally, and Uran Oh. This work is supported by Maryland Procurement Office contract no. H98230-14-C-0137.

8. REFERENCES

[1] Microsoft safety and security center.

[2] US-CERT:Tips.

[3] American community survey 5-year estimates, 2014.

[4] Adams, A., and Sasse, M. A. Users are not the enemy.

[5] Akaike, H. A new look at the statistical model identification.

[6] Akhawe, D., and Felt, A. P. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *USENIX Sec.* (2013).

[7] Arachchilage, N. A. G., and Love, S. A game design framework for avoiding phishing attacks. *Comput. Hum. Behav.* (2013).

[8] Baker, R., Blumberg, S., and et al. AAPOR report on online panels. *The Public Opinion Quarterly* (2010).

[9] Beautement, A., Sasse, M. A., and Wonham, M. The compliance budget: Managing security behaviour in organisations. In *workshop on new security paradigms* (2008).

[10] Bravo-Lillo, C., Komanduri, S., Cranor, L. F., Reeder, R. W., Sleeper, M., Downs, J., and Schechter, S. Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. In *SOUPS* (2013).

[11] Ciampa, M. A comparison of password feedback mechanisms and their impact on password entropy. *Information Management & Computer Security* (2013).

[12] D., B., K., L., and A., A. The Networked Nature of Algorithmic Discrimination. Data and Discrimination: Collected Essays.

[13] Das, S., Kim, T. H., Dabbish, L., and Hong, J. The effect of social influence on security sensitivity. In *SOUPS* (2014).

[14] Das, S., Kramer, A. D., Dabbish, L. A., and Hong, J. I. Increasing security sensitivity with social proof: A large-scale experimental confirmation. In *CCS* (2014).

[15] DeMaio, T. J., Rothgeb, J., and Hess, J. Improving survey quality through pretesting. *U.S. Bureau of the Census* (2003).

[16] Egelman, S., Cranor, L. F., and Hong, J. You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In *CHI* (2008).

[17] Egelman, S., and Peer, E. Scaling the security wall: Developing a security behavior intentions scale (sebis). In *CHI* (2015).

[18] F.R.S., K. P. X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5* (1900).

[19] Fujita, M., Yamada, M., Arimura, S., Ikeya, Y., and Nishigaki, M. An attempt to memorize strong passwords while playing games. In *NBIS* (2015).

[20] Furman, S. M., Theofanos, M. F., Choong, Y.-Y., and Stanton, B. Basing cybersecurity training on user perceptions. *IEEE S&P* (2012).

[21] Furnell, S., Bryant, P., and Phippen, A. Assessing the security perceptions of personal internet users. *Computers & Security* (2007).

[22] Garg, V., Camp, L. J., Connelly, K., and Lorenzen-Huber, L. Risk communication design: Video vs. text. In *PETS* (2012).

[23] Halevi, T., Lewis, J., and Memon, N. A pilot study of cyber security and privacy related behavior and personality traits. In *WWW* (2013).

[24] Hargittai, E. Second-level digital divide: Mapping differences in people's online skills. *First Monday* (2002).

[25] Hargittai, E. *The Digital Divide and What to Do About It.* 2003, pp. 822–841.

[26] Hargittai, E., and Hsieh, Y. P. Succinct survey measures of web-use skills. *Soc. Sci. Comput. Rev.* (2012).

[27] Herley, C. So long, and no thanks for the externalities: The rational rejection of security advice by users. In *NPSW* (2009).

[28] Herley, C. More is not the answer. *IEEE Security & Privacy magazine* (2014).

[29] Hinyard, L. J., and Kreuter, M. W. Using narrative communication as a tool for health behavior change: a conceptual, theoretical, and empirical overview. *Health Educ Behav* (2007).

[30] Hosmer, D. W., and Lemeshow, S. *Applied logistic regression.* 2000.

[31] Howe, A. E., Ray, I., Roberts, M., Urbanska, M., and Byrne, Z. The psychology of security for the home computer user. In *IEEE S&P* (2012).

[32] Hu, X. *Assessing source credibility on social media—An electronic word-of-mouth communication perspective.* PhD thesis, Bowling Green State University, 2015.

[33] Ion, I., Reeder, R., and Consolvo, S. "...no one can hack my mind": Comparing expert and

non-expert security practices. In *SOUPS* (2015).

[34] Ipeirotis, P. Demographics of mechanical turk. *NYU Center for Digital Economy* (2010).

[35] Jerome, J. Buying and Selling Privacy: Big Data's Different Burdens and Benefits. *Stanford Law Review* (2013).

[36] Kang, M. Measuring social media credibility: A study on a measure of blog credibility. *Institute for Public Relations* (2009).

[37] Kang, R., Brown, S., Dabbish, L., and Kiesler, S. Privacy attitudes of mechanical turk workers and the u.s. public. In *SOUPS* (2014).

[38] Kelley, T., and Bertenthal, B. I. Attention and past behavior, not security knowledge, modulate users? decisions to login to insecure websites. *Information and Computer Security* (2016).

[39] Kentaro Toyama, A. R. Kelsa+: Digital literacy for low-income office workers. In *International Conference on Information and Communication Technologies and Development* (2009).

[40] Krosnick, J. A. The threat of satisficing in surveys: the shortcuts respondents take in answering questions. Survey Methods Centre Newsletter, 2000.

[41] Krosnick, J. A. *Handbook of Survey Research*. 2010.

[42] Lin, E., Greenberg, S., Trotter, E., Ma, D., and Aycock, J. Does domain highlighting help people identify phishing sites? In *CHI* (2011).

[43] Mann, H. B., and Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* (1947).

[44] Nappa, A., Johnson, R., Bilge, L., Caballero, J., and Dumitras, T. The attack of the clones: A study of the impact of shared code on vulnerability patching. In *IEEE S&P* (2015).

[45] Poole, E. S., Chetty, M., Morgan, T., Grinter, R. E., and Edwards, W. K. Computer help at home: Methods and motivations for informal technical support. CHI.

[46] Presser, S., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., Rothgeb, J. M., and Singer, E. Methods for testing and evaluating survey questions. *Public Opinion Quarterly* (2004).

[47] Rader, E., and Wash, R. Identifying patterns in informal sources of security information. *J. Cybersecurity* (2015).

[48] Rader, E., Wash, R., and Brooks, B. Stories as informal lessons about security. In *SOUPS* (2012).

[49] Redmiles, E., Malone, A. R., and Mazurek, M. L. How i learned to be secure: Advice sources and selection in digital security. In *IEEE S&P* (2016).

[50] Rice, R. E. Influences, usage, and outcomes of internet health information searching: Multivariate results from the pew surveys. *International J. Medical Informatics* (2006). Health and the Internet for All.

[51] Robila, S. A., and Ragucci, J. W. Don't be a phish: Steps in user education. In *SIGCSE* (2006).

[52] Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who are the crowdworkers?: Shifting demographics in mechanical turk. In *CHI* (2010).

[53] Schaeffer, N. C., and Presser, S. The science of asking questions. *Annual Review of Sociology* (2003).

[54] Schechter, S., and Bonneau, J. Learning assigned secrets for unlocking mobile devices. In *SOUPS* (2015).

[55] Schechter, S. E., Dhamija, R., Ozment, A., and Fischer, I. The Emperor's New Security Indicators. *IEEE S&P* (2007).

[56] Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., and Downs, J. Who falls for phish?: A demographic analysis of phishing susceptibility and effectiveness of interventions. In *CHI* (2010).

[57] Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J., and Nunge, E. Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish. In *SOUPS* (2007).

[58] Siciliano, R. 17 percent of pcs are exposed.

[59] Smith, S. 4 ways to ensure valid responses for your online survey. *Qualtrics*.

[60] Sole, D., and Wilson, D. G. Storytelling in Organizations : The power and traps of using stories to share knowledge in organizations. *Training and Development* (1999).

[61] Stanley, L. D. Beyond access: Psychosocial barriers to computer literacy special issue: Icts and community networking. *The Information Society* (2003).

[62] Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., and Cranor, L. F. Crying wolf: An empirical study of ssl warning effectiveness. In *USENIX Sec.* (2009).

[63] Tourangeau, R., and Yan, T. Sensitive Questions in Surveys. *Psychological Bulletin* (2007).

[64] Ur, B., Kelley, P. G., Komanduri, S., Lee, J., Maass, M., Mazurek, M. L., Passaro, T., Shay, R., Vidas, T., Bauer, L., Christin, N., and Cranor, L. F. How does your password measure up? the effect of strength meters on password creation. In *USENIX Sec.* (2012).

[65] van Dijk, J., and Hacker, K. The digital divide as a complex and dynamic phenomenon. *The Information Society* (2003).

[66] Wash, R. Folk models of home computer security. In *SOUPS* (2010).

[67] Wash, R., and Rader, E. Too much knowledge? security beliefs and protective behaviors among united states internet users. In *SOUPS* (2015).

[68] Whitman, M. E. Enemy at the gate: Threats to information security. *Commun. ACM* (2003).

[69] Willis, G. B. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. 2005.

[70] Wu, M., Miller, R. C., and Garfinkel, S. L. Do security toolbars actually prevent phishing attacks? In *CHI* (2006).

[71] Yuan, M., and Lin, Y. Model selection and estimation in regression with grouped variables. *J. Royal Statistical Society* (2006).