Foundations and Trends® in Privacy and Security

The Security & Privacy Acceptance Framework (SPAF)

A Review of Why Users Accept or Reject Cybersecurity and Privacy Best Practices

Suggested Citation: Sauvik Das, Cori Faklaris, Jason I. Hong and Laura A. Dabbish (2022), "The Security & Privacy Acceptance Framework (SPAF)", Foundations and Trends® in Privacy and Security: Vol. 5, No. 1-2, pp 1–143. DOI: 10.1561/3300000026.

Sauvik Das
Carnegie Mellon University
sauvik@cmu.edu

Cori Faklaris
University of North Carolina
cfaklari@uncc.edu

Jason I. Hong
Carnegie Mellon University
jasonh@cs.cmu.edu

Laura A. Dabbish
Carnegie Mellon University
dabbish@cs.cmu.edu

now
the essence of knowledge
Boston — Delft

Contents

The Security & Privacy Acceptance Framework (SPAF)

Sauvik Das[1], Cori Faklaris[2], Jason I. Hong[1] and Laura A. Dabbish[1]

[1] *Carnegie Mellon University, USA; sauvik@cmu.edu, jasonh@cs.cmu.edu, dabbish@cs.cmu.edu*

[2] *University of North Carolina, Charlotte, USA; cori@corifaklaris.com*

ABSTRACT

How can we encourage end-user acceptance of expert recommended cybersecurity and privacy (S&P) behaviors? We review prior art in human-centered S&P and identified three barriers to end-user acceptance of expert recommendations: (1) awareness: i.e., people may not know of relevant security threats and appropriate mitigation measures; (2) motivation: i.e., people may be unwilling to enact S&P behaviors because, e.g., the perceived costs are too high, and (3) ability; i.e., people may not know when, why, and how to effectively implement S&P behaviors. These three barriers make up what we call the "Security & Privacy Acceptance Framework" (SPAF). We then review and critically analyze prior work that has explored mitigating one or more of the barriers that make up the SPAF. Finally, using the SPAF as a lens, we discuss how the human-centered S&P community might re-orient to encourage widespread end-user acceptance of pro-S&P behaviors by employing integrative approaches that address each one of the awareness, motivation, and ability barriers.

Sauvik Das, Cori Faklaris, Jason I. Hong and Laura A. Dabbish (2022), "The Security & Privacy Acceptance Framework (SPAF)", Foundations and Trends® in Privacy and Security: Vol. 5, No. 1-2, pp 1–143. DOI: 10.1561/3300000026.

1

Introduction

Cybersecurity and privacy (S&P[1]) unlock the full potential of computing. Use of encryption, authentication, and access control, for example, allows employees to correspond with professional colleagues via email with reduced fear of leaking confidential data to competitors or cybercriminals, parents to share photos of children with remote loved ones over the Internet with reduced fear of this data reaching the hands of unknown strangers, and anonymous whistleblowers to share information about problematic practices in the workplace with reduced fear of being outed. Conversely, failure to employ appropriate S&P measures can leave people and organizations vulnerable to a broad range of threats.

In short, the security and privacy decisions we make on a day-to-day basis determine whether the data we share, manipulate, and store online is protected from theft, surveillance, and exploitation. It is unsurprising, therefore, that the compromising of weak security and privacy practices remains the central tenet for a professional cybercrime industry which —

[1]We use the term cybersecurity and privacy to encapsulate the broad concept of protecting digital resources and data from intruders. Cybersecurity is commonly abbreviated to just "security", and so throughout this document we use S&P as shorthand for "cybersecurity and privacy." We use this short-hand in various ways, typically as a descriptor: e.g., S&P threats, S&P behaviors, and S&P tools.

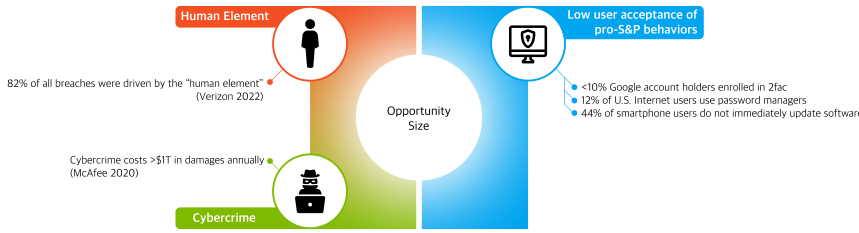

Figure 1.1: Cybercrime is estimated to cause over $1 trillion USD in damages to the global economy, and much of it is enabled by human error. Yet, user acceptance and adoption of expert-recommended security and privacy behaviors remains low. There remains an immense opportunity for impact by improving end-user acceptance and adoption of expert-recommended security and privacy behaviors.

by some estimates — causes upwards of $1 trillion in damages annually to the global economy (Smith and Lostri, 2020).

Many of the data breaches that are responsible for these damages involve human error or manipulation — i.e., improperly configured security settings, the accidental divulsion of key account credentials, or the unwitting installation of destructive malware. Moreover, as an increasing share of economic and social activity is conducted partially or exclusively online, the ramifications of these breaches have never been more significant. In 2021, for example, a ransomware attack crippled the Colonial Pipeline company, causing gas outages all over the eastern seaboard of the United States, resulting in outages, panic and predatory price inflation — and all because the company's private VPN was accessible without multi-factor authentication (Kerner, 2022). The Colonial Pipeline company incident is not an isolated incident. 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 (Moore and Roberts, 2013). In response, stock prices plummeted, adversely affecting thousands. The cause? The AP's Twitter account credentials were phished, and the account was not protected with two-factor authentication. More generally, in 2020, Verizon published an analysis of 3950 security incidents, showing that the most common "actions" that led to breaches were social attacks that prey on human fallibilities (accounting for 22% of all breaches).

Moreover, the authors of that report observed that "the only action type that is consistently increasing year to year in frequency is [human] error." (Verizon, 2020). The 2022 version of that report estimated that the "human element" drove 82% of the 5212 breaches studied (Verizon, 2022). Unsurprisingly, prior work has found that the S&P behaviors that experts recommend only thinly overlap with the behaviors that people find important and adopt (Ion *et al.*, 2015; Busse *et al.*, 2019).

The upshot: if enough people employed basic, expert-recommended best practices — e.g., keeping one's software up-to-date, using multi-factor authentication on important accounts, using a password manager to ensure the reliable use of strong, random passwords unique for each individual account — the cybercrime industry would be hamstrung. The costs of these attacks would be substantially increased, shifting economic incentives, and would likely reduce the prevalence of all but the most sophisticated, targeted attacks. Yet, despite decades of improvements to the usability of S&P systems, end-users still struggle with adopting expert-recommended S&P advice. Indeed, as of early 2018, fewer than 10% of Google account holders had enrolled in two-factor authentication, and at least 17% of Google users reused their account passwords (Milka, 2018). Recent Pew surveys found that only 12% of Internet users in the U.S. use password managers and only 44% immediately update the operating system on their mobile phones (Olmstead and Smith, 2017).

This discrepancy — between the massive damages caused by the exploitation of weak security behaviors, and the existence of security technologies that can significantly reduce these damages, as summarized in Figure 1.1 — begs the question: "How can we encourage end-users to heed the advice of S&P experts?" Put another way, we might ask: "What inhibits acceptance of pro-S&P behaviors among end-users, and how can we overcome those inhibitors?"

In this monograph, we conducted an extensive review of prior literature to answer these questions. We covered a broad range of interdisciplinary perspectives — those from computer science, cognitive, behavioral and social psychology, human-computer interaction, design and behavioral economics. We start with a comprehensive review of extant models of human behavior and technology adoption and use those models as a lens to contextualize prior findings in human-centered

S&P that help explain why end-users accept or reject pro-S&P behaviors (see Section 2).

We found that there are three key inhibitory barriers to pro-S&P behaviors: awareness, motivation, and ability (see Section 3). First, many consumers are unaware of S&P threats that may be pertinent to a given situation, nor the techniques and tools that can be used to counteract these threats. Second, many consumers are unwilling to employ the techniques and tools that are available to protect against common threats. Third, many consumers are unable to correctly use the techniques and tools that are available to protect against common threats. Taken together, this triplet of inhibitory barriers make up what we call the "Security and Privacy Acceptance Framework" (SPAF). Efforts to address one or more of these inhibitory barriers can be said to increase acceptance of expert-recommended (pro-)S&P behaviors; efforts that — intentionally or not — exacerbate these barriers can be said to decrease acceptance of pro-S&P behaviors.

We next reviewed the existing body of work in human-centered S&P aimed at increasing end-user acceptance of pro-S&P behaviors (see Section 4) — particularly in the usable privacy and security, behavioral economics, human-computer interaction, and social psychology domains. Using the SPAF as a lens, we then critically analyzed why, despite decades of improvements to the usability of end-user S&P systems, widespread acceptance of pro-S&P behaviors remains relatively low (see Section 5). Specifically, we argue that while many existing interventions have been shown to be effective at addressing one or more of the barriers in the SPAF, there are relatively few interventions that target all barriers at once. Integrative approaches that target awareness, motivation, and ability at once are likely to be more effective at driving end-user acceptance and adoption of pro-S&P behaviors. We conclude by synthesizing promising trends and directions for future work (also Section 5).

A final note: in this monograph, we primarily focus on encouraging S&P behaviors that protect users against third-party and interpersonal threats, often making the assumption that a first-party service provider can be trusted. We acknowledge that security and privacy enhancing technologies can also be used to protect oneself against first-party and

institutional threats, but argue that protection against these threats is less straightforward from the perspective of end-user action — indeed, placing the onus strictly on end-users is a problematic approach. For these situations, there may be a stronger need for regulation of bad-faith corporate and intelligence agency practices, rather than targeted design interventions and behavioral improvements on the part of end-users.

2

Background

In order to uncover why end-users struggle with adopting and practicing expert-recommend S&P behaviors, a necessary first step is to understand factors that drive human behavior and technology adoption more generally. Indeed, models of human behavior can afford insight into what drives general decision making, while models of technology adoption offer illustrative insights into what factors explain the spread of technologies among populations of interest. From that baseline understanding, we then identify factors that are unique to S&P technologies and behaviors more specifically. Note that our goal here is not to exhaustively review all models of human behavior and technology adoption, but to ground prior literature on why people accept or reject expert-recommended S&P behaviors in the broader literature on human behavior. A summary of theories of human behavior and technology adoption we reviewed can be found in Table 2.1.

Table 2.1: Summary of the theories of human behavior and technology adoption that we analyzed as they relate to explaining why users accept or reject expert-recommended security & privacy advice.

Theory	Core Insight
Models of Human Behavior	
Theory of Reasoned Action (Fishbein, 1979)	Behavioral intention is a function of attitudes and subjective norms. In S&P, attitudes often conflict with behavior.
Theory of Planned Behavior (Fishbein, 1979)	Adds perceived behavioral control to the theory of reasoned action. Users only act if they feel like their actions matter.
Fogg Behavioral Model (Fogg, 2009)	Behavior is a function of motivation, ability, and trigger. Persuasive design should focus on enhancing motivation and ability, or catalyzing action.
Computer-Human Information Processing model (Wogalter, 2006a)	Highlights the cognitive barriers that influence risk-mitigating behaviors. These barriers pertain to channel of delivery, attention, comprehension & memory, beliefs & attitudes, and motivation.
Models of Technology Adoption	
Diffusion of Innovations (Rogers, 1962)	Codifies the processes through which an innovation spreads through members of a social system. This diffusion is affected by five factors: relative advantage, compatibility, complexity, trialability, and observability.
Technology Acceptance Model (Davis, 1989)	Identifies three non-independent factors that affect individual user acceptance of information technologies in organizational contexts. The factors include perceived ease of use, perceived usefulness, and behavioral intention. Subsequent versions of the model include antecedents to these factors.
Unified Theory of Acceptance and Use of Information Technology (Venkatesh *et al.*, 2003)	Synthesizes information systems research on technology adoption into one unified model of technology adoption that comprises four factors: performance expectancy, effort expectancy, social influence, and facilitating conditions. Also identifies four moderating variables: age, gender, experience, and voluntariness.

2.1 Models of human behavior

An explanatory model of human behavior is one of the holy grails of research in psychology. In our review, we consider primarily models that can be used to predict either behavioral intention (Fishbein and Ajzen, 1977) — an individual's perceived likelihood or subjective probability that they will engage in a given target behavior — or the target behavior itself.

2.1.1 The theories of reasoned action (TRA) and planned behavior (TPB)

A popular, early model of human behavior is the theory of reasoned action (Fishbein, 1979). The theory of reasoned action posits that behavioral intention — which is the immediate antecedent to a target behavior — is a function of an individual's attitude towards a target behavior and the subjective norms an individual associates with that behavior (Fishbein, 1979). In other words, an individual is more likely to engage in a target behavior if they believe that the target behavior is likely to result in a specific desired outcome and if they believe that doing so will be perceived positively, or not negatively, by others. In the context of security and privacy behaviors, the marked difference between self-reported attitudes and observed behavior in end-user S&P have led some to postulate the existence of a "privacy paradox" (Norberg *et al.*, 2007): the idea that despite people claiming to desire the properties of S&P in their use of computing systems, they have low behavioral intention for following through on expert-recommended S&P advice. The TRA offers an explanation — if attitudes disagree with behavioral intention, then perhaps the subjective norms people associate with following through with S&P behaviors can explain the difference. As we shall see, the emerging discipline of social cybersecurity helps model the effects of subjective norms in S&P decision making — in particular, prior work suggests that the early adopters of S&P tools can be perceived by others as paranoid which, in turn, can inhibit adoption of expert-recommended S&P behaviors (Das *et al.*, 2014a; 2015).

The theory of reasoned action, however, is predicated under the assumption of full volitional control: i.e., the assumption that an individual believes they are in full control of the outcomes that follow their action (Fishbein, 1979). The theory of planned behavior complicates this model by introducing a new antecedent to both behavioral intention and behavior: perceived behavioral control. Perceived behavioral control reflects the extent to which a subject believes that their behavior alone might result in a desired outcome. The higher the perceived behavioral control, the higher the behavioral intention and the likelihood of the target behavior. In the context of S&P, evidence from prior work suggests that many end-users tend to have low perceived behavioral control. For example, prior work has shown that users express concern, anger, and frustration when they encounter institutional privacy violations — be it through investigative journalism, as in the Snowden revelations (Landau, 2013), or through personal exposure to data breaches, like the Equifax breach (Wikipedia, 2021). Yet, a 2019 Pew study found that over 80% of adults in the U.S. believed that they had little or no control over the data that corporations and the government collected, and that it was impossible to go through daily life without having data about themselves collected (Auxier *et al.*, 2019). Other strands of work highlight that some users believe that is not their responsibility to keep their data secure; rather, it is the responsibility of the service provider (Carre *et al.*, 2018; Haney *et al.*, 2021).

2.1.2 The Fogg Behavioral Model (FBM)

While the TRA and TPB offer helpful theoretical lenses to explain human behavior, they were developed as descriptive models that do not necessarily offer design implications. A popular prescriptive model of human behavior is the Fogg Behavior Model (FBM) (Fogg, 2009). According to the FBM, behavior occurs if and only if an individual wants to adopt the behavior (i.e., has motivation), is easily able to adopt the behavior (i.e., has the ability) and something prompts action (i.e., something triggers that specific behavior) (Fogg, 2009). Accordingly, to design persuasive technologies that encourage specific target behaviors, the FBM suggests that one must aim to affect either a user's motivation or ability, or deliver appropriate triggers.

In the context of S&P, motivation and ability are both well understood "barriers" to pro-S&P behaviors. The field of usable privacy and security traces its origins to identifying and addressing the ability challenges in user-facing security systems (Zurko and Simon, 1996; Whitten and Tygar, 1999; Adams and Sasse, 1999). As early as 1996, for example, Zurko and Simon discuss the need for "user-centered security" in which they outline a vision for "considering user needs as a primary goal at the start of secure system development." (Zurko and Simon, 1996) In their seminal 1999 paper, "Why Johnny Can't Encrypt", Whitten and Tygar systematically uncovered a wide array of usability flaws that inhibited average end-users from properly using PGP 5.0 to encrypt email correspondences (Whitten and Tygar, 1999). Around that same time, Adams and Sasse offered a rebuttal against the prevailing notion of "users being the weakest link" in security by arguing that security lacked a user-centered design process which, in turn, resulted in security controls that were fundamentally unusable (Adams and Sasse, 1999). Unsurprisingly, since these canonical contributions, there have been many proposed design interventions and technologies that aim to address the ability barrier — e.g., by making user-facing security technologies faster and more intuitive (we will explore much of this work in Section 4).

Motivation in S&P, likewise, is well studied. In studying security practices "in-the-wild", Dourish *et al.* (2004) found that security is a "secondary concern" for end-users: i.e., while security is an attribute that most of us would claim to want in our use of computing systems, it is peripheral to our primary task at any given moment (e.g., checking email, managing our finances). Given the secondary nature of security concerns, therefore, it is unsurprising that users have low motivation to handle security-relevant interruptions (e.g., updating their software in the middle of a workday) and that few users are proactive in their approach to security (Das *et al.*, 2019a). From the perspective of behavioral economics, Herley (2009) argued that the expected costs — to end-users — of following all security advice they might receive could out weight the benefits of following that advice. Redmiles *et al.* (2018) later reported on a series of experiments to show many users do, in fact, behave rationally when making security decisions — weighing costs as they relate to

risks. Gaw *et al.* (2006) and Das *et al.* (2014a) illustrated how social influences might negatively impact users' motivation to be secure—use of weighty security solutions can sometimes be considered "paranoid" and this perception can, in turn, inhibit motivation to use those security solutions. We provide more complete coverage of factors that impact end-user motivation in Section 3.

Triggers are direct antecedents to specific behaviors — e.g., the warning that alerts the user to financial fraud, the notification that reminds a user to update their software. Fogg (2009) defines three types of behavioral triggers for persuasive design: *sparks*, which motivate people with high ability but low motivation; *facilitators*, which simplify action for people with high motivation but low ability; and, *signals*, which serve as reminders for people who already have high motivation and ability. Many existing S&P warnings and notifications are signals. Sparks and facilitators also pose interesting opportunities for S&P, as few end-users have both high motivation and high ability to engage in pro-S&P behaviors. An example of a spark that encourages S&P behaviors is Das *et al.* (2014b)'s social proof notifications, which informed Facebook users of the number of their friends who used optional security tools on Facebook. An example of an effective facilitator that simplifies S&P behaviors comes from Akhawe and Felt (2013)'s redesign of the Chrome SSL warning to simplify exiting out of suspicious webpages. Through an online survey with 852 users on Amazon Mechanical Turk, Das *et al.* (2019a) introduced a typology of perceived direct antecedent triggers for S&P behaviors—making a distinction between social, proactive and forced triggers. Social triggers came from other individuals, be it in the form of direct observation, conversations or stories. Proactive triggers came from within the individual — an internally motivated decision to take action. Forced triggers required participants to take action — e.g., an employer or service mandating a password change or enrolling in two-factor authentication. The authors found that social triggers were the most commonly reported direct-antecedent behavioral triggers for people with low-moderate security behavioral intention, forced triggers were the most commonly reported for people with low security behavioral intention, and proactive triggers were the most commonly reported for people with high security behavioral intention (Das *et al.*, 2019a).

2.1.3 Wogalter's C-HIP

The TRA, TPB and FBM are models of general human behavior, but there are also models of risk-mitigating behaviors in particular. The Computer-Human Information Processing model (C-HIP), introduced and developed by Wogalter (2006b), is one such model and helps explain how consumer product warnings and other cautionary triggers influence risk-mitigating behavior in particular. The C-HIP is a stage model that helps explain the factors and cognitive barriers that play a in role in impacting end-user behavior that results from the perception and/or delivery of a warning. Specifically, these factors and steps include: source, or the entity that is transmitting a warning; channel, or the mechanism(s) through which the warning is delivered; attention switch, or the transitioning of one's attention to the warning from something else; attention maintenance, or the sustained attention on a warning to parse its messaging; comprehension & memory, or the understanding of gleaned information such that one knows what is the risk and what to do about it, and the encoding of this information into memory; beliefs & attitudes, or people's pre-existing orientation towards the product for which a warning is being delivered or the information of the warning itself; motivation, or factors that influence whether or not action is carried out after a user is aware of, comprehends and believes in a warning; and, behavior, or compliance with the target behavior that is being sought by a warning.

The C-HIP mirrors the FBM in a number of ways: the source and channel are factors that belong to trigger in the FBM; attention switch, attention maintenance, and comprehension & memory are factors that translate to awareness in the FBM; and, beliefs & attitudes and motivation are factors that translate to motivation in the FBM. The C-HIP, however, provides specific prescriptive guidance on designing S&P warnings and interruptions to encourage risk mitigating behavior. For example, owing to the effects of habituation — i.e., the diminished individual effects of warnings upon repeated exposure to similar warnings (Vance *et al.*, 2017; 2018) — it is important to design S&P warnings that are salient and distinctive. Prior research in usable security has observed this effect both empirically and physiologically, e.g., through

an analysis of fMRI scans of people's brains upon being presented with a warning both initially and after many repeated exposures (Vance *et al.*, 2017; 2018). Prior research has also explored interventions that can be used to counteract this effect — e.g., through the use of interactive elements that command end-user attention and comprehension before dismissal is possible (Bravo-Lillo *et al.*, 2013).

Moreover, Wogalter (2006b) argues that warnings should only be designed for "hidden hazards" — i.e., hazards that are not readily apparent. Knives, for example, are sharp but almost every adult knows that they are sharp and thus no warning is needed to alert people to their sharpness. In contrast, the door to a room with invisible, noxious fumes would warrant a warning. Unlike the physical world, however, very little is intuitively apparent in the context of cybersecurity — indeed, nigh every cybersecurity hazard can be considered "hidden" or "abstract" when one considers the fact that threats are often remote and asynchronous.

2.2 Models of technology adoption, diffusion, acceptance

Outside general models of human behaviors, scholars from sociology, marketing, and STS have studied and modeled technology adoption behaviors specifically.

2.2.1 Rogers' Diffusion of Innovations (DoI)

One canonical model in this space is Rogers' Diffusion of Innovations (DoI) (Rogers, 1962). Rogers defines the DoI as the process through which an innovation — a technology, solution, or product that is perceived as "new" — is communicated through certain channels over time among the members of a social system. Rogers goes on to identify five factors that produce the degree to which we can expect innovations to diffuse in a given social system. These five factors are: *relative advantage*: the degree to which an innovation is perceived as better than the idea it supersedes; *compatibility*: the degree to which an innovation is perceived as being consistent with the values, experiences and needs of potential adopters; *complexity*: the degree the which an innovation is perceived as

difficult to use; *trialability*: the degree to which an innovation may be experimented with on a limited basis; and, *observability*: the degree to which the results of an innovation are visible to others (Rogers, 1962).

Rogers' five factors, too, are highly correlated with the FBM. Indeed, relative advantage and compatibility are factors that affect an individual's motivation; complexity and trialability affect ability; and, observability affects awareness. However, where Fogg's behavior model might help explain an individual's behavior, Rogers' DoI model helps explain how behaviors spread in a population.

The DoI also posits that the potential pool of adopters can be categorized into one of five categories depending on how early they are likely to adopt a new technology: innovators (the earliest 2.5%), early adopters (13.5%), early majority (34%), late majority (34%) and laggards (16%). The relative importance of the five aforementioned factors varies across these groups, and thus strategies to promote the adoption of novel innovations should vary depending on its existing market saturation (Rogers, 1962).

While the DoI is an informative model for all new technologies, in later work, Rogers introduces a new class of innovations that diffuse differently: "preventive" innovations (Rogers, 2002). Preventive innovations require action at one point in time in order to avoid an unwanted future condition — medication, for example, or insurance. S&P innovations are preventive innovations.

Preventive innovations diffuse more slowly than other types of innovations. Rogers argues that the relative advantage of an innovation is the most important predictor of how quickly it is likely to diffuse. Yet, because the benefits of preventive innovations are reaped at a future time (or never at all), their perceived relative advantage is almost always low (Rogers, 2002). Empirical work seems to confirm this hypothesis in the context of S&P — e.g., in discussing why people appear to shirk strong privacy settings on social media, Acquisiti describes the concept of hyperbolic time discounting, or the tendency for people to over-value the immediate gratification of posting over the future reward of stronger privacy (Acquisti and Grossklags, 2005; Acquisti *et al.*, 2017). Given that S&P innovations are meant to prevent an undesirable future condition (i.e., a S&P breach), and given that their success may be

completely invisible (because if they work properly, the user observers no noticeable change in their condition), the DoI suggests that S&P innovations are doomed to diffuse slowly.

To overcome this inertia, Rogers argues that it is imperative to take advantage of interpersonal diffusion channels—i.e., peer-to-peer recommendations. As we will see, however, today's S&P tools are broadly designed in a manner that is individualistic and fails to activate peer networks.

2.2.2 The Technology Acceptance Model (TAM)

Whereas Rogers' DoI considers the spread of innovations, in general, Davis' Technology Acceptance Model (TAM) identifies factors that affect user acceptance of information technologies in organizational contexts (Davis, 1989). The TAM, proposed in its original form by Fred Davis, originates in the theories of reasoned action and planned behavior — Davis argued that neither one could reliably predict user acceptance of a novel information system. The TAM posits that there are three non-independent factors that explain an individual's acceptance of a novel information technology: (i) perceived ease of use; (ii) perceived usefulness; and, (iii) behavioral intention. Perceived ease of use and perceived usefulness are considered to be antecedents to behavioral intention that, in turn, can be affected by design. Note that these factors are not unlike Rogers' five factors — indeed, perceived ease of use can be likened to Rogers' complexity factor, while perceived usefulness can be likened to Rogers' relative advantage factor.

TAM is one of the most widely cited and influential models of information technology adoption. Unsurprisingly, then, in the decades that have passed since it was first proposed, there have been a number of significant updates resulting in both a TAM 2 (Venkatesh and Davis, 2000) and a TAM 3 (Venkatesh and Bala, 2008). TAM 2 complicates the perceived usefulness and perceived ease of use variables and introduces measurable antecedents that might affect those variables. Five factors are considered to be antecedents to perceived usefulness: *subjective norms*, or the influence of others on the user's decision to use or not to use the technology; *image*, or the desire of the user to maintain a

favorable standing among others; *job relevance*, or the degree to which the technology was applicable; *output quality*, or the extent to which the technology adequately performed the required tasks; and *result demonstrability*, or the production of tangible results. Two broad factors affect perceived ease of use: *anchors*, or general beliefs about computers and computer usage; and, *adjustments*, or beliefs that are shaped by direct experience. Subjective norms, perceived behavioral control and user specific characteristics all affect anchors, while trialability, visibility, result demonstrability and content richness all affect adjustments. TAM3 (Venkatesh and Bala, 2008) posits that direct experience could also moderate computer anxiety, perceived ease of use and, thus, behavioral intention. While TAM has been used in disparate research contexts since its inception, the empirical evidence in support of these models have come primarily from organizational contexts.

The information systems community has also proposed a unified theory of acceptance and use of information technology (UTAUT) (Venkatesh *et al.*, 2003). Like the TAM, TRA and TPB, behavioral intention is seen as the direct antecedent to technology adoption. In turn, the UTAUT identifies four key factors (i.e., performance expectancy, effort expectancy, social influence, and facilitating conditions) and four moderators (i.e., age, gender, experience, and voluntariness) that are useful in predicting behavioral intention, synthesizing insights from a broad range of competing models. Performance expectancy is defined as the degree to which an individual believes that using a system will help with one's job performance (like relative advantage in the DoI and perceived usefulness in TAM); effort expectancy is the degree of ease associated with the use of the system (like complexity in the DoI and perceived ease of use in TAM); social influence is the degree to which an individual perceives that important others believe they should use the new system (related to observability and compatibility in DoI and subjective norms in TAM); and, facilitating conditions are the degree to which an individual believes that the technical infrastructure exists to support use of the system (related to trialability in DoI and perceived ease of use in TAM). Given that the UTAUT, itself, was broadly inspired by a number of the theoretical models of human behavior and technology acceptance previously covered, many of these factors and their consideration in security and privacy may seem familiar.

TAM and the UTAUT have been broadly influential in organizational contexts in predicting and facilitating the spread of novel information systems. However, the applicability of these models is less well understood in the context of preventive innovations (e.g., S&P) and in non-organizational contexts (e.g., for home computer users). To date, most applications of TAM and UTAUT in the context of S&P have been proposed extensions to the original models that introduce privacy and security concerns as distinct factors that can influence the adoption of information systems more generally (e.g., Roca *et al.*, 2009). There have been no concrete attempts, that we know of, to adapt the TAM or UTAUT to predict the adoption of S&P behaviors specifically.

3

The Security & Privacy Acceptance Framework

3.1 Why do we need a framework specific to S&P acceptance?

Based on a synthesis of the extant literature in usable privacy and security and scaffolded by the aforementioned theories of human behavior and technology adoption, we propose a framework to help explain the acceptance and diffusion of expert-recommended S&P behaviors — the security and privacy acceptance framework (SPAF). But what is the need for a framework specifically calibrated to cybersecurity behaviors?

For the past two decades, research from the usable privacy and security community has uncovered at least three ways in which S&P behaviors are distinct from general behavior and pose unique challenges in encouraging widespread acceptance (see also Figure 3.1):

- First, many S&P behaviors are *preventive*. Like health behaviors, S&P tools and best practices help people avoid an undesirable future state (e.g., one in which their personal data is compromised to a cybercriminal or an intelligence agency) that may never, in fact, come to pass. As Rogers argues, preventive behaviors have inherently lower perceived relative advantage than other behaviors, limiting their adoption and spread (Rogers, 2002).

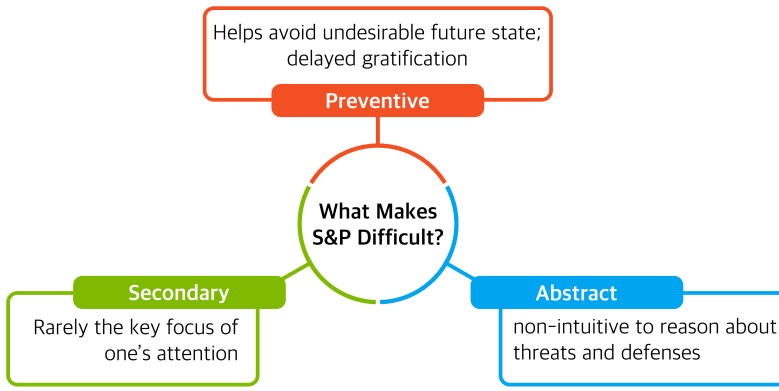

Figure 3.1: Cybersecurity and privacy behaviors are difficult in a way that is distinct from general behaviors in at least three ways — they are preventive, secondary, and abstract. Taken together, existing frameworks of human behavior and technology adoption are not as explanatory for S&P behaviors.

- Second, S&P behaviors are *secondary*. While most would claim to desire the property of security in their interactions with computing systems, the process of ensuring one's S&P is always peripheral to one's primary goal or task at any given moment (which might be, e.g., checking email, preparing a document). The secondary nature of security and privacy behaviors was noted by Dourish *et al.* (2004) in their study of end-user security behaviors in the wild.

- Third, the mechanisms through which S&P behaviors protect one's data are *abstract*. When we push a heavy object against a door, we have an intuitive understanding that any attempt to open that door will be met with strong resistance. When we update our software, few of us have an intuitive understanding why or how doing so makes us more secure. More generally, the human mind is not attuned to the broad invisibility, ubiquity and remoteness of S&P threats; in turn, the actions experts prescribe to protect against these threats do not always inspire urgency or concern. Whitten and Tygar (1999) discussed the abstraction property as a key inhibitory barrier for general adoption of S&P tools.

Because S&P behaviors are preventive, secondary and abstract, general models of human behavior and technology adoption are not as predictive in the context of S&P: the rules that govern whether a person, for example, elects to start an exercise regimen, learn a new language, or download a new word processor either do not apply or must be heavily modified to predict whether an individual will, for example, commit to keeping their software updated, or enable the use of two-factor authentication for important accounts, or use a VPN when on public WiFi access points.

In the broadest strokes, the SPAF consists of three factors that are strongly inspired by extant models of human behavior — particularly the FBM (Fogg, 2009) and C-HIP (Wogalter, 2006b). The three factors can be considered non-independent and necessary prerequisites to S&P acceptance: awareness, motivation, and ability (see also Figure 3.2). The sub-factors that make up these factors, in turn, are strongly inspired by the factors described in the Diffusion of Innovation (Rogers, 1962; 2002), TAM (Davis, 1989; Venkatesh and Davis, 2000; Venkatesh and Bala, 2008), and UTAUT (Venkatesh *et al.*, 2003).

Figure 3.2: The Security and Privacy Awareness Framework (SPAF) comprises of three barriers that must be overcome to encourage end-user acceptance of expert-recommended S&P behaviors: awarness, motivation, and knowledge. Prior work suggests that four factors influence user S&P awareness: social engagement, mental models, media exposure, and warnings & notifications. Four factors also affect user S&P motivation: subjective norms, perceived relative advantage, trialability, and compatability. Finally, two factors affect user S&P ability: system usability / complexity and accessibility.

Awareness encompasses whether or not an individual understands threats that are material to the data, accounts, and devices they would like protected, and the expert-recommended mitigation measures to protect against those threats. Factors that can affect awareness include social engagement, mental models and digital literacy, media exposure,

training, and warnings / notifications. People who are unaware of a threat cannot take measures to mitigate 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 — even if they are motivated to do so and are able.

Motivation encompasses whether or not an individual wants to employ expert-recommended best practices and tools to protect their data, devices, and accounts. Factors that affect motivation include subjective norms, perceived relative advantage, trialability, and compatibility. Individuals — even those who are aware of threats and able to protect themselves against those threats — may not take protective or preventative action if they believe that doing so might be perceived as "paranoid", if they believe S&P is not their responsibility, or if they assess the costs of taking expert-recommended S&P action to be too high relative to the perceived advantages, for example.

Ability encompasses whether or not an individual is capable of converting intention into action by utilizing expert-recommended best practices and tools to protect their data, devices and accounts. People — even those who are aware and motivated — may be unable to appropriately put expert-recommended S&P advice into effect if the tools they use make doing so difficult, if expert-recommended advice is rife with jargon, or if they have never been formally trained. Factors that affect ability include how well the tool addresses Norman (2013)'s usability gulfs of execution and evaluation, as well as the accessibility of a system to users with disabilities.

In the sections that follow, we provide a more detailed overview of the extant empirical work that demonstrates how any one of these factors can serve as an inhibitory barrier to the acceptance and diffusion of expert-recommended S&P behaviors.

3.2 Awareness

The FBM posits that specific behaviors are catalyzed by triggers — a warning, a notification, a conversation, an observation (Fogg, 2009). The C-HIP was specifically calibrated to a very specific type of trigger: a warning (Wogalter, 2006b). In both models, behavior starts with acute

awareness of threats and preventive actions that are brought about by a catalyzing trigger. In both models, however, individuals' behaviors and the triggers that catalyze them are considered in isolation.

For preventive and secondary behaviors, like S&P, people often require repeated exposures to catalyzing triggers before taking action. Health interventions, for example, are often not adopted after just one exposure but many (Kaplan *et al.*, 2012). The direct antecedent trigger, thus, can be thought of as the proverbial straw that breaks the camel's back — the last in a long string of exposures and influences that has raised users' awareness about why a specific behavior or action might be necessary. The construct of awareness in the SPAF, then, is better understood as a person's broader understanding of why a S&P behavior might be necessary, what it protects, and what threats it protects against, rather than the more acute awareness brought about by a warning or interruption.

Many users have low awareness of security threats and the tools available to protect themselves against those threats. For example, Adams and Sasse (1999) found that insufficient awareness of security issues caused users to construct their own model of security threats that were often incorrect, ultimately resulting in security breaches. Stanton *et al.* (2004) found that a lack of awareness of foundational security principles even led some "experts" to make security mistakes, such as using a social security number as a password. More generally, in the context of S&P, prior work suggests that there are at least four factors that affect an individual's awareness of both material S&P threats as well as the extant expert-recommended defenses against those threats: social engagement, mental models and digital literacy, media exposure, and warnings & notifications.

The factors listed above are not all mutually independent and have, in fact, demonstrated interaction effects. For example, Das *et al.* (2019a) conducted a survey with over 800 people from the U.S. and India, inquiring about the direct-antecedents to recent S&P behaviors. They found that users with low security behavioral intention were most likely to have engaged in a specific S&P behavior because they were forced to do by, e.g., a mandatory system update or employer policy, while those with low-medium security behavioral intention were mostly likely to do

so because of social engagement by, e.g., observing a peer or having a conversation. Those with the highest security behavioral intention were more likely to simply engage in expert-recommended S&P behaviors proactively, suggesting high inherent awareness that required no external catalyst (Das *et al.*, 2019a).

3.2.1 Social engagement

Social engagement encompasses how interactions with other people, either active or passive, influences end-user awareness of both S&P threats and pro-S&P behaviors. In their analysis of the direct antecedents to pro-S&P behaviors, Das *et al.* (2019a) found that social triggers were the most commonly reported, particularly for users with mid-range security behavioral intention. In quantitatively analyzing user-reported sources of security advice and behavior, Redmiles *et al.* (2016a) also found that social information sources can impact awareness: in particular, individuals with lower digital literacy and education tended to rely more on the advice of family and friends when making security and privacy decisions.

DiGioia and Dourish alluded to the need for social awareness as early as 2004 when introducing the concept of "social navigation" for security and privacy behaviors (DiGioia and Dourish, 2005). Like a worn path through an otherwise well manicured lawn, the authors argue, social navigation systems that show end-users traces of what other users do when they encounter similar security and privacy decisions can help raise awareness of both security and privacy threats and appropriate responses to those threats.

In a qualitative analysis of how users make security decisions, Rader *et al.* (2012) argued that "security stories" that people hear from others can strongly influence what threats users find pertinent. Das *et al.* (2014a) also found that conversations between people — typically in the form of cautionary tales or social sensemaking of potential threats — influenced people's awareness of security and privacy threats. They also found that the observability of S&P tools that can be used to protect against common threats (e.g., Android's 9-dot pattern lock) could impact people's awareness of S&P tools (Das *et al.*, 2014a). In a

later experiment with 50,000 Facebook users, Das *et al.* (2014b) found that increasing the observability of the use of optional security and privacy tools was significantly more likely to result in end-users exploring the adoption of those tools themselves. Prior work has also found that social sharing of pertinent media information can increase awareness of S&P threats and the mitigation strategies thereof (Das *et al.*, 2018b).

3.2.2 Mental models and digital literacy

End-user mental models of digital technology and their general digital literacy can also impact their awareness of security and privacy threats. Mental models "describe how a user thinks about a problem" (Wash, 2010), whereas digital literacy more broadly encompasses the cognitive, technical and socio-emotional skills associated with the skillful use of digital technologies (Ng, 2012).

In a qualitative analysis of how end-user mental models of the Internet, Kang *et al.* (2015) found that people with more articulated technical models of the Internet perceived more privacy threats than those with more vague understandings of the Internet. Non-experts, for example, focused on third-party "hackers" and other people as threats, whereas experts also thought about institutional threats (e.g., the corporations and governments that controlled the underlying Internet infrastructure).

Wash (2010) also found that folk models of security threats tended to correlate with the defensive strategies users were aware of and found pertinent. For example, users had different perceptions of security threats and the defensive strategies those threats entailed: some believed that threats primarily arose from young mischief makers, where as others believed that threats resulted from professional criminals.

Likewise, Yao *et al.* (2017) analyzed people's folk models of online behavioral advertising and found that both technically savvy and non-technically savvy people tended to have inaccurate or incomplete models that, in turn, affected their overall attitudes towards and perceptions of S&P threats.

Mismatched mental models can also be indicative of low digital literacy: i.e., knowledge about how information technologies work and

can be used. Redmiles *et al.* (2016a) found that individuals with low digital literacy tend to become aware of cybersecurity threats through a different set of information sources than do those with higher digital literacy. Specifically, they rely more on family and friends, service providers and media exposure than do those with higher digital literacy.

3.2.3 Media exposure

Prior work has also found that security and privacy news can drive awareness of both threats and mitigation strategies, though how people hear about these events and how much they rely on these events to become aware of relevant security and privacy threats can vary based on personal characteristics such as age and security behavioral intention.

In a qualitative investigation, Das *et al.* (2014a) found that news articles were the most frequent catalysts for security and privacy related conversations, for example; these conversations, in turn, could lead to social sensemaking processes in which multiple individuals collaboratively determined the importance of a threat and how to respond to it. Redmiles *et al.* (2016b) also found that the media was one of the most prevalent sources of security and privacy advice to which individuals are exposed. Specifically, they found that the media was the primary source of advice for threats and behavioral best practices associated with passwords and two-factor authentication.

In a subsequent quantitative analysis of how S&P news events spread, Das *et al.* (2018b) found that different people came to learn about security and privacy news events through different sources — younger people, for example, relied more on online news sources, social media and friends, whereas older people relied more on broadcast and television. Nicholson *et al.* (2019), in exploring the cybersecurity information seeking behaviors of older adults, also found that older adults relied strongly on broadcast sources to assess whether or not a given security threat was important.

3.2.4 Warnings & notifications

Finally, another factor that affects end-user awareness of S&P threats and the tools that can be used to protect against those threats are

warnings (indications of imminent danger or risk) and notifications (indications of pertinent security and privacy information that do not necessarily entail immediate danger). For example, notifications to update software, interruptions that alert users to the presence and status of a website's SSL certificate, and permission prompts can all be considered security and privacy relevant warnings and notifications that aim to raise users' awareness. In their study of the direct antecedent triggers that preceded S&P behaviors, Das *et al.* (2019a) categorized warnings and other such external interruptions as "forced" triggers and found that people with low security behavioral intention were most likely to engage in an expert-recommended S&P behavior as a result of such warnings. Based on a long history of research on warning design in usable privacy and security we now understand that the design of warnings can significantly impact their effectiveness at commanding end-user attention and adherence (see, e.g., Bravo-Lillo *et al.*, 2013; Akhawe and Felt, 2013; Egelman *et al.*, 2008b; Anderson *et al.*, 2015), and that different designs are necessary for different contexts (Schaub *et al.*, 2015).

Yet, one of the most pernicious problems for raising awareness of security and privacy threats through external triggers such as warnings is habituation — the gradually reduced effectiveness of a trigger at commanding users' attention (Wogalter, 2006a). Prior work has found, for example, that even well-designed security and privacy warnings have reduced effectiveness as users are increasingly exposed to those warnings. Moreover, habituation to stimuli has deep-seated neurological origins in the human brain. For example, Vance *et al.* (2017, 2018) ran a study in which they repeatedly exposed participants to cybersecurity and privacy warnings, and collected fMRI brain scans of participants on each exposure. They found a general decline of participants' attention to warnings on subsequent exposures and that participants adherence to static warnings decreased dramatically over three weeks.

This habituation response may be rational, as users are confronted with many warnings that are not associated with a real threat. Indeed, as Bravo-Lillo *et al.* (2013) note: "today's computer systems perpetually cry for attention in the name of safety, and hundreds of cries may be heard without a real threat". There have been a number of proposals

designed to mitigate the habituation effect: Anderson *et al.* (2015) found that polymorphic warnings that change in appearance are more resilient to habituation effects, Bravo-Lillo *et al.* (2013) found that introducing interactive elements into a warning that users must engage with prior to clicking-through on an option can command end-user attention, Egelman *et al.* (2008b) found that active warnings that interrupt users are more likely to be heeded than passive warnings, and Akhawe and Felt (2013) found that the visual and interaction design of a warning can significantly improve users adherence even on repeated exposures. However, the habituation effect remains present even if reduced. Indeed, as Sunshine *et al.* (2009) note, "while warnings can be improved, a better approach may be to minimize the use of warnings....altogether."

In the next section, we shall explore in more detail the methods proposed to improve the effectiveness of warnings at commanding user attention and reducing the habituation effect.

3.3 Motivation

Awareness, alone, does not result in action: people must also be willing and motivated to act. However, users are often unmotivated to behave in accordance with expert-recommended S&P recommendations — an observation that dates back to the origins of the usable security and privacy field of study (Zurko and Simon, 1996; Whitten and Tygar, 1999; Adams and Sasse, 1999). Inspired by the factors that make up general models of human behavior and technology adoption, our review of prior work suggests that there are 4 factors that impact end-user motivation towards accepting and/or adopting expert-recommended security advice: subjective norms, perceived relative advantage, trialability, and compatibility.

3.3.1 Subjective norms

Subjective norms are perceived expectations from others that influence a user to perform a particular behavior (Ajzen, 1991). Subjective norms have long been understood to be key drivers of human behaviors — indeed, subjective norms play an essential role in driving behavioral

intentions in both the theory of reasoned action and the theory of planned behavior (Ajzen, 1991; Fishbein and Ajzen, 1977) as well as in technology adoption in both the DoI (Rogers, 1962), the technology acceptance model (TAM) (Davis, 1989), and the universal theory of acceptance and use of technology (UTAUT) (Venkatesh *et al.*, 2003). In the context of security and privacy, prior work suggests that subjective norms and social influences can strongly influence people's motivation to accept and/or adopt expert-recommended security and privacy advice (Rader *et al.*, 2012; Das *et al.*, 2014a; 2014b; 2019a).

In early work in this space, DiGoia and Dourish postulated that "social navigation" — or, increasing the observability of other users' S&P actions in a computing system — would likely motivate other users to emulate similar behaviors (DiGioia and Dourish, 2005). The authors argued that people navigate complicated systems and spaces in the physical world based partially on their understanding and interpretations of the activities of others, yet security and privacy features are often left to users to navigate themselves. Rader *et al.* (2012) found that informal stories from other users motivated specific risk mitigating behaviors such as, e.g., "not sharing personal information online" — though these behaviors were not necessarily in alignment with expert recommendations. Das and colleagues found that social proof — the implicit knowledge or direct observation of others engaging in security and privacy behaviors — was a key motivator of security and privacy behavior changes (Das *et al.*, 2014a; 2014b; 2019a). In an initial qualitative study, a number of users reported enabling Android 9-dot pattern lock because they saw other users doing the same (Das *et al.*, 2014b). In a subsequent survey study, Das *et al.* (2019a) found that social triggers were the most commonly reported direct antecedents to security and privacy relevant behavioral changes. Finally, in a randomized, controlled experiment with 50,000 Facebook users, Das *et al.* (2014b) found evidence to suggest that notifications that contained information about how many of one's friends used optional security features were significantly more likely to motivate viewers to explore those features for their own use than a notification that did not contain this social information. Krsek *et al.* (2022) further found that social proof derived from the public-at-large was as effective as social proof from friends and experts at influencing pro-S&P behavior in the Facebook context.

Stanton *et al.* (2004) explored motivational antecedents to security behaviors in organizational contexts and found that organizational culture was a key motivator for positive security behaviors. In particular, workers in environments in which information security was more universally valued — i.e., in which security concern was more of a subjective norm — were more likely to engage in positive security behaviors.

It is also important to note that subjective norms can have negative effects on the adoption and acceptance of pro-S&P behaviors, as well. Specifically, if users perceive that others around them get by without accepting expert-recommended security and privacy advice, they may be less likely to follow the advice themselves. Prior work exploring the efficacy, on strong password composition, of a password meter that compared the strength of a user's password to that of others found evidence of a possible "boomerang" effect in which users with initially strong passwords reduced the strength of their final passwords to more closely approximate the norm (Egelman *et al.*, 2013). In a quantitative analysis of how friends' adoption of optional security tools on Facebook affects one's own adoption of those same tools, Das *et al.* (2015) found that individuals with only a few friends who used those optional systems were less likely to adopt those systems themselves.

3.3.2 Perceived relative advantage

Rogers (1962) defines relative advantage as the "degree to which an innovation is perceived to be better than the idea it supersedes." Moreover, Rogers (2002) argues, perceived relative advantage is the *most important* predictor of the rate of adoption of innovations. In the context of security and privacy, a core motivational challenge is that unless there are mandates from an authority (e.g., enforced use of two-factor authentication by employees of a company), the baseline that a security and privacy innovation must "supersede" is not using such a technology at all. Given that security and privacy technologies commonly introduce new interaction barriers (e.g., an additional login step with two-factor authentication, slower data retrieval with encrypted drives) for invisible, abstract, and future-oriented benefits, the perceived relative advantage of an S&P innovation is likely to be low.

Security and privacy are what Dourish *et al.* (2004) call secondary concerns. Many people want their digital data, resources and interactions online to be securely stored and transmitted, but they are not using digital services to be secure — rather, they may be trying to communicate important professional information to work colleagues, or sharing photos of young children with loved ones in a different country, or streaming a video to learn how to fix their car. It is difficult to motivate end-users to engage in specific behaviors aligned with primary concerns (e.g., updating one's Word processor with improved productivity features); it is significantly more difficult to motivate end-users to engage in behaviors aligned with secondary concerns. Indeed, Rogers (2002) argues that preventive innovations — those that help prevent undesirable future outcomes, which are typically secondary concerns peripheral to a user's core objective at any given moment — are perceived as having lower relative advantage in general.

Prior work suggests that that many security threats remain abstract to many individuals (Adams and Sasse, 1999; Herley and Oorschot, 2009; Whitten and Tygar, 1999): e.g., Bob may know, conceptually, that there are security risks to using the same simple password across accounts, but may not believe that he is, himself, in danger of experiencing a security breach. Additionally, Herley (2009) 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. 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. Beautement *et al.* (2008)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. Moreover, because the perceived benefits of security and privacy tools are delayed to some unknowable time in the future, these

benefits may be subject to "hyperbolical discounting" — i.e., perceived as having less value than they do given that they are delayed (Acquisti and Grossklags, 2005). More generally, users often reject the use of security and privacy tools when they expect or experience them to be weighty (Adams and Sasse, 1999; Sasse, 2003; Gaw *et al.*, 2006).

Another factor that affects perceived relative advantage is expectancy and/or perceived agency, or a user's belief that the actions they take can materially affect desired outcomes. Prior work has found that users sometimes feel helpless with respect to S&P, believing that if an attacker wanted to access their data they would irrespective of any counter-measures taken (Wash, 2010; Spiekermann, 2007). This learned helplessness is particularly true of perceptions of institutional surveillance threats (Auxier *et al.*, 2019). First-party and third-party personal data harvesters increasingly employ sophisticated tracking and profiling technologies to create and monetize detailed digital portraits of users at-scale (Zuboff, 2015). In the face of this global tracking apparatus, a 2019 Pew study found that over 80% of adults in the U.S. believed that they had little or no control over the data that corporations and the government collected, and that it was impossible to go through daily life without having data about themselves collected (Auxier *et al.*, 2019). In turn, this learned helplessness can reduce the perceived relative advantage of employing protective technologies and advice.

3.3.3 Trialability

Trialability is the "degree to which an innovation may be experimented with on a limited basis" (Rogers, 1962). While some security expert-recommended security advice can be trialed, the combination of the immediacy of the costs of a security technology combined with the invisibility and abstractness of security threats makes it difficult to assess the potential benefits of a security technology through a trial, but easy to assess the immediate costs. Herley (2016) argues that security claims are often unfalsifiable — one can never make any guarantees that their data is definitively "secure", only that they have not yet been compromised.

Other expert-recommended security and privacy advice cannot be trialed at all: e.g., keeping one's software up-to-date. Prior work suggests that negative experiences with software updates can inhibit one's desire to keep one's systems and softwares updated in the future (Vaniea *et al.*, 2014). In a survey of over 200 security experts and non-experts, Ion *et al.* (2015) found that while experts valued keeping software up-to-date, non-experts reported being skeptical of the effectiveness of these behaviors because of prior negative experiences with updates. Moreover, when security updates are bundled with other updates, negative experiences with unrelated updates — e.g., user interface changes — can also impact users' motivations to install future security-relevant updates (Wash *et al.*, 2014).

3.3.4 Compatibility

Rogers (1962) defines compatibility as the "degree to which an innovation is perceived as being consistent with the existing values, past experiences and needs of potential adopters." In the context of security and privacy, compatibility can be thought of as how well expert-recommended security advice fits with users' workflows, identities and perceptions.

Prior work suggests that security and privacy innovations, in general, score low on compatibility for most users. Indeed, security measures are often antagonistic towards the specific goal of the end user at any given moment (Dourish *et al.*, 2004; Sasse, 2003). 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 (Egelman *et al.*, 2010). Prior work also suggests that some users may not always view securing their digital resources, accounts and data as their own responsibility (Haney *et al.*, 2021) — they may view it as the responsibility of the service provider.

Social influences can also impact compatibility — specifically, users' conceptions of how compliance with security and or privacy advice might be perceived by others. Gaw *et al.* (2006), for example, found that people perceive weighty security and privacy strategies to secure

email communications as "paranoid". Das *et al.* (2014b), in exploring how social influences impact end-user security and privacy behavior changes, found that expert users avoid sharing their knowledge with non-expert friends to avoid coming off as "preachy" or "nutty". In a subsequent quantitative analysis in partnership with Facebook, Das *et al.* (2015) found that when users are connected to only a few other users who use optional security tools like two-factor authentication on Facebook, the presence of those few friends appeared to negatively correlate with a user's own likelihood to adopt those optional security tools. In explaining this effect, Das (2017) introduced the "parnaoia-disaffiliation" hypothesis: the postulation that because early-adopters of S&P tools can be perceived by others as paranoid, general users reject use of these tools because they develop an illusory correlation between use of security features and being paranoid. Indeed, not needing security because one has "nothing to hide" is a common misconception that inhibits the uptake of security and privacy advice (Solove, 2007).

3.4 Ability

Even users who are aware of a threat and motivated to act may not know how to appropriately act to protect themselves against those threats. In other words, users often have trouble converting intention into action. Prior work suggests that two broad factors affect ability: system usability and accessibility / inclusivity.

3.4.1 System usability

Using Don Norman's usability gulfs as an illustrative lens, we categorize prior work demonstrating ability challenges in security and privacy according to two gulfs: the gulf of execution and the gulf of evaluation (Norman, 2013). The gulf of execution is the degree to which the interaction possibilities of a system correspond with a user's intentions and what the user perceives as possible (Norman, 2013). In the context of security and privacy, the gulf of execution is small when one knows how to utilize a given system to protect their data, devices and resources from pertinent threats. The gulf of evaluation is the degree to which the

system provides feedback and representations that can be interpreted in terms of the expectations and intentions of the user (Norman, 2013). In the context of security and privacy, we can consider the gulf of evaluation to be small when one can tell how their security and privacy decisions affect the protection of their data, devices and resources against pertinent threats. Identifying and addressing the gulfs of execution and evaluation in consumer-facing security and privacy systems has long been acknowledged as a central tenet of the usable privacy and security community.

The gulf of execution

Prior work suggests that there is a wide gulf of execution for most security features for most users. A canonical example that helped lay the foundation for the usable privacy and security community comes from Whitten and Tygar (1999). The authors critically analyzed the user interface for PGP 5.0 — an email client designed to allow end-users to send encrypted emails. Using a cognitive walkthrough methodology with cryptography novices, the authors found significant usability challenges including, for example, information overload, confusing metaphors (e.g., 'public key' and 'private key'), understanding the practical intention and purpose of cryptographic operations such as signing messages. Around the same time, Adams and Sasse published another article that is now part of the usable security and privacy canon: "Users Are Not the Enemy" (Adams and Sasse, 1999). In that paper, the authors conducted an online questionnaire through which they analyzed user perceptions and behaviors relating to password systems in two organizations. Their analysis uncovered that poor end-user security behaviors are likely the result of a failure to employ human-centered design processes in designing password systems, resulting in systems that require users to engage in practices antagonistic to their workflows and that assume in-depth knowledge of computer security.

The usable privacy and security community has since expanded on these analyses. Indeed, there have been at least two replications of Whitten and Tygar (1999) for modern encrypted email systems: e.g., PGP9 (Sheng *et al.*, 2006) and Mailvelope (Ruoti *et al.*, 2015). The

upshot of this prior work is that 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. For example, many users cannot distinguish legitimate vs. fraudulent URLs, nor forged vs. legitimate email headers (Dhamija *et al.*, 2006). Another study revealed how security features in Windows XP, Internet Explorer, Outlook Express, and Word applications are difficult for lay users to navigate (Furnell *et al.*, 2006). Ion *et al.* (2015) found that the privacy and security behaviors of non-experts and experts only thinly overlap, owing largely to a knowledge gap between the two groups. And, almost every year, we see new critical analyses of security and privacy systems and controls that are unusable: e.g., FIDO2 roaming authenticators (Owens *et al.*, 2021).

In general, security systems must be usable, fast, and comprehensible by non-experts if they are to minimize the gulf of execution.

The gulf of evaluation

In the context of security and privacy, the gulf of evaluation entails whether or not a user understands how their behaviors and decisions materially impact the protection of their data, devices and resources against pertinent threats. However, the abstractness and invisibility of security threats combined with the unfalsifiability of security claims (Herley, 2009) makes bridging the gulf of evaluation a longstanding challenge.

First, prior work suggests that security and privacy concerns, and the expert-recommended advice to protect against these concerns, are difficult for end-users to evaluate because they are abstract (Whitten and Tygar, 1999) — attackers are remote and invisible, may or may not exist, will only be material at some unknowable point in the future, and the mechanisms through which they may attack are difficult for non-experts to comprehend. Indeed, in an analysis of users' mental models of how the Internet works and how attackers might compromise one's personal data in transmission, Kang *et al.* (2015) found that non-expert end-users had simple and vague models relative to the more technical and articulated models of experts. These more articulated models, in turn, exposed more avenues through which attackers might

intercept and edit personal data. Wash (2010) similarly found that home computer users harbored folk models of security threats that were sometimes simplistic and that attackers could exploit gaps in these users' knowledge to circumvent their defensive efforts.

Moreover, Herley (2009) argued that many security claims are unfalsifiable — it is not possible, for example, to definitely state that one is "secure" only that one has not yet been compromised. This unfalsifiability, in turn, reduces users' ability to assess whether or not the expert-recommended security advice they are given will help protect against pertinent threats and for how long. For example, systems that employ adversarial machine learning to allow users to evade facial recognition surveillance have recently been proposed (Shan *et al.*, 2020; Cherepanova *et al.*, 2021; Chandrasekaran *et al.*, 2020), but the efficacy of these systems are in constant flux as new defenses against these adversarial techniques are proposed. While an inevitable byproduct of the arms race inherent to security technologies, even expert end-users can have difficulty understanding to what extent they are protected, for how long, and against whom.

3.4.2 Accessibility / Inclusivity

Beyond usability, the accessibility of S&P technologies can also exacerbate or mitigate ability barriers for users. Users are not a monolith; for example, they may have different physical abilities, come from different cultural contexts, have different digital literacies, and have differing levels of educational attainment. These differences, in turn, can materially impact from where users get S&P advice (Redmiles *et al.*, 2016a), how disruptive or demanding a given S&P practice might be (Bigham and Cavender, 2009; Ma *et al.*, 2013), and whether a user is capable of independently performing a particular behavior (Murthy *et al.*, 2021). Recognizing the broad diversity of people and how some populations of users are under-served by existing S&P controls, Wang (2018) outlined a vision for "inclusive security" that is "concerned with designing security and privacy mechanisms that are inclusive to people with various characteristics, abilities, needs, and values." Note that our coverage of the relevant literature here is necessarily limited given the breadth and diversity of the human experience — our objective here is simply

to demonstrate that differences across user sub-populations materially impact ability barriers in end-user S&P.

As an example, while S&P technologies are often inhibiting in nature, the encumbrances they impose are sometimes disproportionately worse for people with disabilities. CAPTCHAs, for example, are a nuisance for most users but are one of the most significant accessibility hurdles for people with visual impairments (WebAIM, 2017). Even alternatives designed to be more accessible — like audioCAPTCHAs — are slower and more cognitively demanding for people with visual impairments than visual CAPTCHAs are for others (Bigham and Cavender, 2009). Cognitive disability can also exacerbate ability barriers for S&P technologies and practices. In a comparative analysis of graphical and alphanumeric password systems among neurotypical users and users with Down syndrome, Ma *et al.* (2013) found that those with Down syndrome found the creation and use of graphical passwords to be significantly more demanding and less preferable. One caveat, however, is that is important not to over-generalize from one population to another. For example, Marne *et al.* (2017) found that some graphical passwords were actually easy to learn and usable for participants with learning disabilities such as Dyslexia and Dyscalculia. In other words, interventions that address ability barriers for one population may inadvertently introduce ability barriers for other populations.

Digital literacy can also create ability barriers. Some older adults, may have low digital literacy (Hargittai and Dobransky, 2017), which can cause them to be disproportionately targeted by Internet fraud (Reisig *et al.*, 2015) and complicate their ability to enact expert-recommended S&P behaviors (Frik *et al.*, 2019b). As a result, prior research has demonstrated that users with lower literacy often call upon social connections with higher literacy to act as S&P stewards or caregivers (Murthy *et al.*, 2021; Kropczynski *et al.*, 2021), though this ad-hoc practice is rarely inherently supported by existing S&P controls and can result in stewards having too much power.

In sum, one-size-fits-all approaches cannot adequately address the broad diversity of people who use computing technologies and, therefore, must interact with S&P controls to keep their digital accounts, devices, and resources protected.

3.5 Summary

Using extant models of human behavior and technology adoption as a lens, we synthesized prior literature spanning usable privacy and security to introduce the security and privacy acceptance framework (SPAF). Note that the SPAF, in and of itself, is not an empirically validated model but a framework through which we can systematically reason about the factors that underlie users' decisions to accept or reject expert-recommended security and privacy advice. The SPAF posits that there are three high-level factors that ultimately explain what leads to end-user acceptance and adoption of expert-recommended S&P advice: awareness, motivation and ability. Awareness encompasses a user's knowledge of threats pertinent to the data, devices and resources they would like to be protected, as well as the expert-recommended advice and strategies to protect against those threats. Motivation encompasses factors that underlie end-users' willingness to follow expert-recommended advice to protect their data, devices and resources against pertinent threats. Finally, ability encompasses factors that encompass whether or not a user is capable of converting intention into action by accurately enacting expert-recommended security and privacy advice. We next review the interventions, systems and tools designed to encourage more widespread adoption of expert-recommended security and privacy advice as they relate to the SPAF.

4

Encouraging Widespread Security & Privacy Acceptance

The SPAF systematizes factors that lead to user acceptance or rejection of expert-recommended security and privacy advice, but can also help categorize attempts to improve acceptance of expert-recommended security advice. We next review attempts to improve end-user awareness of, motivation to adopt, and ability to implement expert-recommended security and privacy advice. As we will discuss in the sections to follow, efforts have been made at improving all parts of the SPAF stack. Nevertheless, it is important to note that 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 other interventions to increase end-user acceptance of pro-S&P behaviors. Moreover, while we have attempted to categorize prior art at improving awareness, motivation and ability in the subsections to follow, the most effective interventions address, at least partially, multiple different barriers in the SPAF — i.e., these interventions are not necessarily exclusive to addressing just one barrier. Table 4.1 summarizes the prior work we surveyed as they relate to addressing the barriers in the SPAF.

Table 4.1: Prior work on encouraging end-user acceptance of S&P categorized as addressing awareness, motivation, and/or ability. The majority of prior work has focused on addressing the ability barrier, though there have also been attempts at addressing the awareness and motivation barriers. While some prior art may at least peripherally address two barriers, we found no prior work that has empirically validated addressing all three.

⊕ *Work that has empirically validated addressing the barrier in question.*
⊙ *Work that hypothesizes addressing the barrier in question.*

	Addresses Awareness?	**Addresses Motivation?**	**Addresses Ability?**
Awareness campaigns			
Khan *et al.* (2011)	⊙	⊙	
Kajzer *et al.* (2014)	⊕		
Strand (2018)	⊕		
Bada *et al.* (2019)	⊕	⊙	
Scrimgeour and Ophoff (2019)	⊕		
Simulated Attacks			
Ferguson (2005)	⊕		
Dodge Jr *et al.* (2007)	⊕		
Jansson and Solms (2013)	⊕		⊙
Kumaraguru *et al.* (2008)	⊕		⊙
Kumaraguru *et al.* (2009)	⊕		⊙
Warnings / Notifications			
Egelman *et al.* (2008b)	⊕		
Kelley *et al.* (2009)	⊕		
Bravo-Lillo *et al.* (2013)	⊕	⊕	
Felt *et al.* (2015)	⊕	⊙	
Anderson *et al.* (2015)	⊕		
Wilson *et al.* (2017)	⊙		
Petelka *et al.* (2019)	⊕		⊙
Napoli *et al.* (2020)	⊙		
Do *et al.* (2021a)	⊕	⊙	
Games			
Sheng *et al.* (2006)	⊕		⊕
Denning *et al.* (2013)	⊙	⊙	
Dabrowski *et al.* (2015)	⊙	⊕	
Alotaibi *et al.* (2017)	⊙		⊙
CJ *et al.* (2018)	⊕		
Mostafa and Faragallah (2019)	⊕		
Alqahtani and Kavakli-Thorne (2020)	⊙		
Chen *et al.* (2020)	⊙	⊕	
Nudges			
Wang *et al.* (2014)	⊙	⊕	
Almuhimedi *et al.* (2015)	⊕	⊕	
Nicholson *et al.* (2018)		⊕	
Frik *et al.* (2019a)		⊕	
Story *et al.* (2020)	⊙	⊕	
Golla *et al.* (2021)		⊕	
Pro-social design			
Goldschlag *et al.* (1999)		⊙	⊙
DiGioia and Dourish (2005)	⊙	⊙	
Goecks *et al.* (2009)		⊙	⊕
Bonneau *et al.* (2009)			⊙
Das *et al.* (2017)		⊙	⊕
Egelman *et al.* (2013)		⊙	
Maxwell (2013)		⊙	⊙
Benet (2014)			⊙
Das *et al.* (2014b)		⊕	⊕

Table 4.1: Continued.

	Addresses Awareness?	Addresses Motivation?	Addresses Ability?
Ohyama and Kanaoka (2015)		⊕	
Dupuis and Khan (2018)		⊕	⊕
Mendel *et al.* (2021)			⊕
Krsek *et al.* (2022)		⊕	
Logas *et al.* (2022)		⊙	⊕
Wu *et al.* (2022a)		⊙	⊕
Akter *et al.* (2022)	⊙	⊕	
Zhang *et al.* (2021)			⊕
Improving passwords			
McCarney *et al.* (2012)			⊕
Bonneau and Schechter (2014)			⊕
Blocki *et al.* (2014)			⊕
Stobert and Biddle (2014)		⊙	⊙
Das *et al.* (2016)			⊕
Mayer *et al.* (2016)			⊙
Yıldırım and Mackie (2019)	⊙		⊕
Shay *et al.* (2015)		⊙	⊕
Ur *et al.* (2017)			⊕
Guan *et al.* (2017)		⊙	⊙
Stobert *et al.* (2020)			⊕
Password alternatives			
Monrose and Rubin (1997)			⊕
Mantyjarvi *et al.* (2005)			⊕
Brainard *et al.* (2006)			⊙
Alsulaiman and El Saddik (2006)			⊙
Alsulaiman and El Saddik (2008)			⊙
Wobbrock (2009)			⊕
Schechter *et al.* (2009)			⊕
Thorpe *et al.* (2013)			⊕
Hayashi *et al.* (2013)			⊕
Das *et al.* (2013)			⊕
Hang *et al.* (2015)			⊕
Karapanos *et al.* (2015)			⊕
Woo *et al.* (2016)			⊕
George *et al.* (2019)			⊕
Das *et al.* (2019b)			⊕
Usable access control			
Kapadia *et al.* (2004)			⊕
Bauer *et al.* (2007)		⊙	⊕
Egelman *et al.* (2008a)			⊕
Reeder *et al.* (2008)			⊕
Wang *et al.* (2009)			⊙
Reeder *et al.* (2011)			⊕
Mazurek *et al.* (2011)		⊙	⊕
Vaniea *et al.* (2012)			⊕
Mazurek *et al.* (2014)			⊙
Klemperer *et al.* (2012)			⊕
Secure messaging			
Ruoti *et al.* (2016)			⊕
Lerner *et al.* (2017)			⊕
Do *et al.* (2021c)	⊙		⊙
Accessibility / inclusivity			
Bigham and Cavender (2009)			⊕
Azenkot *et al.* (2012)			⊕
Barbosa *et al.* (2016)			⊕
Jain *et al.* (2019)			⊕
Havron *et al.* (2019)	⊙		⊙
Fanelle *et al.* (2020)			⊕

Table 4.1: Continued.

	Addresses Awareness?	Addresses Motivation?	Addresses Ability?
Intelligent automation			
Sadeh *et al.* (2009)			⊙
Zhao *et al.* (2014)			⊙
Liu *et al.* (2016b)			⊙
Liu *et al.* (2016a)			⊕
Do *et al.* (2021b)	⊙		⊕

4.1 Improving Awareness

Prior work seeking to improve end-users' awareness of S&P threats fall into three broad categories: awareness training campaigns, warning & notifications, and tools to improve operational / situational awareness of threats.

4.1.1 Awareness training campaigns

Awareness campaigns are time-bound strategies to increase visibility, among either a targeted population or the public-at-large, of general S&P threats and the pertinent expert recommendations to combat these threats. Awareness campaigns have been explored, for example, to improve information security awareness among employees in organizations (e.g., Lebek *et al.*, 2013), among broader populations of users in multi-national contexts (e.g., Bada *et al.*, 2019), and among end-users of specific systems (e.g., Das *et al.*, 2014b). The specific tactics and strategies used in awareness campaigns can vary. For example, Facebook annually shows a subset of its users a notification informing them of the presence of optional-use security tools like two-factor authentication and login notifications (Das *et al.*, 2014b). Other awareness campaigns may involve presentations or training that employees at a company are required or encouraged to view.

Prior literature suggests that awareness campaigns can modestly improve end-user awareness of cyber-threats. Indeed, in conducting a review of the efficacy of awareness campaigns at improving information security awareness in institutions of higher learning, Rezgui and Marks (2008) found that institutions that have awareness campaigns were considered more successful and advanced in information security than

institutions without. Scrimgeour and Ophoff (2019) found that an information security awareness campaign in an organization in South Africa resulted in a "slight increase in the user's knowledge of security controls in their day-to-day working environment."

However, prior work also suggests awareness campaigns often fail to result in actual behavior change, and that the messaging, cultural context, and the personalities and attitudes of those to whom the campaigns are presented can all impact efficacy. Siponen (2000) used the theory of planned behavior and the technology acceptance model as a lens to critically analyze how awareness campaigns fall short, finding that awareness campaigns that contain only descriptive messaging of threats are unlikely to change behavior and that what is needed is prescriptive information that more directly corresponds to specific behavioral changes. In a more recent meta-analysis of why awareness campaigns often fail to change end-user behavior, Bada *et al.* (2019) argued that awareness campaigns are likely insufficient if the broader goal is to change behavior — awareness, the authors argue, is just one factor that leads to behavior change. Indeed, much prior work on awareness campaigns do conflate awareness of behavior with behavioral intention. As we argue in presenting the SPAF in the previous section, however, awareness is just one of three factors that are necessary for the acceptance and adoption of pro-S&P behaviors. Bada *et al.* (2019) also compared awareness campaigns in the U.K. and Africa and found that the messaging was culturally situated: campaigns in the UK, for example, emphasize individuals and the need to protect one's own data, whereas campaigns in Africa appear to be more likely to refer to fulfilling one's duties and obligations to a social group and/or community.

Beyond situating the messaging of an awareness campaign, Kajzer *et al.* (2014) suggest that the messaging in awareness campaigns should be tailored to individual personality types. In evaluating how participants who varied across a range of personality psychometrics — i.e., Big Five personality traits, Machiavellianism, and social desirability — responded to awareness messages framed in terms of deterrence, morality, regret, feedback, and incentive, the authors found some modest effects. Specifically, people who scored higher on agreeableness and neuroticism were more likely to respond to awareness messages framed in terms of

deterrence, whereas people who scored higher on social desirability were less likely to respond to awareness messages framed in terms of morality (Kajzer *et al.*, 2014). In a mixed methods analysis of how a security awareness campaign deployed in an organization of 2000 knowledge workers was differently received by different employees, Strand (2018) found differences in "gender, age and management responsibility level with regard to information security in the workplace." The upshot of this prior work is that awareness campaigns may need to be tailored to individuals to be maximally effective.

Social proof and normative information might also motivate positive responses to messaging campaigns. In an analysis of messaging used in health awareness campaigns, Khan *et al.* (2011) posit that the use of normative signals — i.e., information about one's peers behaviors and attitudes — may result in improved compliance with and efficacy of security awareness campaigns. One possible explanation for this effect is that including normative information can also be a motivational boost, addressing both awareness and motivation at once. In organizational contexts, Strand (2018) also found that management and peers can play an influential role in how security awareness campaigns are received.

4.1.2 Simulated phishing attacks

Phishing encompasses a broad range of social engineering attacks in which an attacker attempts to "fraudulently acquire sensitive information from a victim by impersonating a trustworthy third party" (Jagatic *et al.*, 2007) — for example, account credentials, credit card numbers, personally identifiable information. Phishing remains one of the most pernicious end-user facing problems in cybersecurity today. Unsurprisingly, then, improving end-user awareness of phishing threats remains of strong academic and industrial interest.

In organizational and institutional settings, simulated phishing attacks are commonly utilized to improve member / employee awareness of phishing threats. Tracing its root to military security exercises, in simulated phishing attacks, organizations spoof innocuous emails that *should* be identified as phish as a means of measuring how many of their members might fall prey (Ferguson, 2005; Dodge Jr *et al.*, 2007). Some

of these simulated phishing attacks are further embedded with training materials designed to educate victims about phishing threats (e.g., on identifying common hallmarks of phish) (e.g., Jansson and Solms, 2013) — the idea being that the moment during which an individual falls prey to a simulated phish is a "teachable moment" in which educational interventions are likely to be particularly effective (Kumaraguru *et al.*, 2008).

Prior work has illustrated that simulated phishing attacks, particularly those with embedded training, are effective at raising awareness of phish. Moreover, Kumaraguru *et al.* (2009) found that the effects of these phishing interventions can last at least as long as 28 days and that simulated phishing attacks do not appear to decrease users' willingness to click on links in legitimate messages.

While simulated attacks have been shown to raise awareness of phishing attacks in organizational contexts, there is comparatively less data on the effectiveness of such simulated attacks for security awareness outside of phishing and outside of organizational contexts. In addition, the ethics of running simulated phishing attacks outside of employee security training and assessment has also come under scrutiny. For example, Jagatic *et al.* (2007) ran a simulated spear phishing attack on university students by scraping publicly accessible social meta-data from students' social media profiles to personalize the phishing messages being sent, but the study was met with significant backlash by those who were targeted.

4.1.3 Warning & notification design

Warnings and notifications are commonly used to inform end-users about security and privacy threats and/or tools, advice and behaviors that can help protect users from those threats. These warnings and notifications can be both reactive (shown in response to an action initiated by the end-user) or proactive (shown independent of user action or context to encourage general awareness). There is ample prior work on improving end-user awareness of cybersecurity threats through improved warning and notification design in many distinct domains. Note, however, that while warnings and notifications primarily improve

awareness they can also improve motivation and ability — Fogg (2009) notes that some behavioral triggers might be "sparks" (that improve motivation) while others might be "facilitators" (that simplify desired action). In the context of improving S&P acceptance, however, much of the extant work has focused on improving awareness: generally by assessing how well a warning captures users' attention when competing with other tasks and information, as well as by reducing habituation — the tendency for users to more weakly react to warnings that are similar to ones they have seen in the past (Vance *et al.*, 2018).

A dominant thrust of prior work on improving S&P warnings and notifications focuses on improving interaction, visual and information design. For example, Egelman *et al.* (2008b) found that, in web browsers, active warnings that commanded a user's attention were more likely to be noticed than passive warnings. However, active interruptions disrupt end-user experience and should be avoided unless critical. Felt *et al.* (2014; 2015) improved end-user adherence to Google Chrome SSL warnings by reducing technical jargon, improving visual appeal and carefully designing choice architectures to make safe options the default action. Bravo-Lillo *et al.* (2013) introduced the idea of inhibitive attractors — interactive, animated elements that attract users attention to pertinent information when it seems like they might make a hazardous choice (e.g., clicking through a warning), and found that users were more likely to make an informed decision when exposed to these attractors. Anderson *et al.* (2015) found that polymorphic warnings — those that are visually varied upon repeated exposures to the same user — are more resilient to habituation effects. Kelley *et al.* (2009) explored redesigning privacy policies as "nutrition labels", finding that their approach allowed participants to more quickly and accurately find pertinent privacy information as compared to a plain text baseline.

Finally, the positioning and context of presentation also matters. Petelka *et al.* (2019) explored and experimentally evaluated warnings to reduce end-user susceptibility to email phish. Specifically, they tested positioning phishing warnings close to suspicious links in emails, displaying warnings on hover interactions over links in emails, and by forcing users to click only on raw URLs instead of aliased hyperlinks. In a between-subjects experiment with 701 users, the authors found that

placing warnings near email links and forcing users to pay attention to raw URLs were the most effective at reducing users' susceptibility to email phish. In a broad review of privacy notices, Schaub *et al.* (2015) synthesized a design space for effective privacy notices to help designers reason about factors that are important to consider when designing new privacy notifications.

Prior work has also explored non-visual warning modalities. For example, Micallef *et al.* (2017) explored combinations of modalities for smartphone notifications and found that people preferred privacy notifications to be non-auditory and distinct from other types of notifications. Beyond user preference, Vance *et al.* (2018) found neurological evidence that non-essential notifications can blur with security warnings when distributed through the same communication channels, further motivating the need to communicate security and privacy information through distinct notification channels. Wilson *et al.* (2017) introduced thermal feedback as a new way to communicate web browser security warnings to a user by examining the association between people's perception about temperature and the level of security. Furthermore, Napoli *et al.* (2020) discuss that the effect that thermal stimulation has on an end-user's cognition can help improve their security awareness. Their preliminary findings suggest that thermal stimulation could help communicate the security of TLS certificates to end-users. However, they also note that this thermal feedback might confuse end-users because the heating pad used to deliver the thermal warning may contain residual heat after users navigate away from an insecure website. Do *et al.* (2021a) designed and evaluated Spidey Sense, a wristband that produces customizable squeezing sensations to alert users to urgent cybersecurity warnings. Through a series of lab studies, they found that "squeeze" notifications were judged by end-users as being more appropriate for important and urgent cybersecurity warnings than a vibrotactile buzz.

Finally, services like "Have I Been Pwned?"[1] and password health services offered by password managers like Dashlane help improve end-users operational awareness of when their data may have been included in data breaches. However, the specific effects of these monitoring

[1] https://haveibeenpwned.com/

services on end-user awareness and behavior have yet to be empirically measured.

4.1.4 Games

Games have been broadly explored as a mechanism to improve end-user, employee and student awareness of key S&P concepts and threats. While games can be used to improve motivation and ability as well, they have canonically been explored in the context of improving S&P awareness. Indeed, in a review of a number of games developed for cybersecurity education and training, Hill Jr *et al.* (2020) found that many existing such games can be categorized into one of four categories — security awareness, network and web security, cryptography, and secure software development — and that, of these categories, games developed to improve security awareness were the most populous. In particular, prior literature has explored the use of "serious" and/or "transformational" games — i.e., games designed to change the knowledge, attitudes and/or behaviors players even after playing the game — to improve cybersecurity awareness and knowledge.

Games for improving employee awareness of organizational cyber threats

Many of the original games developed in the context of S&P education and training were developed to improve employee S&P awareness in organizational contexts. In a broader review of 28 papers, Hendrix *et al.* (2016) attempted to assess whether games were generally suitable for employee S&P awareness training. While they found that most of the papers they reviewed lacked a rigorous evaluation, the 11 papers that did include an empirical evaluation of games on S&P awareness were generally positive. One of the earliest and most influential examples of games being used to improve employee cybersecurity awareness is Sheng *et al.* (2007)'s Anti-Phishing Phil, in which end-users were taught to distinguish between authentic and phishing websites based on attributes of the URL of that website. Through a between-subjects experiment with a general population of Internet users, the authors found that end-users who played Anti-Phishing Phil versus receiving an informational

packet were better able to identify fraudulent websites, suggesting that games can be effective in improving awareness. In a similar vein, CJ *et al.* (2018) recently designed and evaluated a similar narrative-style phishing awareness training game, Phishy. Through an evaluation with 8071 employees in an enterprise setting, the authors found that such games were effective at improving phishing awareness for individuals who were novices or intermediate at cybersecurity and privacy, but had little effect on experts.

Games for teaching cybersecurity concepts to students

Games have also been explored as a means of teaching S&P concepts to students. For example, Mostafa and Faragallah (2019) developed six games to loosely approximate games from prior work, all in different genres, and evaluated the pedagogical effectiveness of these games at improving students' S&P awareness with 81 undergraduates in a controlled experiment. Treatment groups used the games the authors had developed, while the control group used lecture notes. The dependent variable that was compared between the two groups was students' performance on a written test. The authors found that students in the treatment groups performed significantly better on the written test than those in the control group. However, there also remains significant room for improvement in educational S&P game design. In a review of 183 publications related to cybersecurity games/gamification for educational purposes, Roepke and Schroeder (2019) found that existing games primarily teach factual knowledge but note that "to actually teach more sustainable knowledge or skills in CS, we need to teach a mixture of factual, conceptual and procedural knowledge".

Games for improving the general public's awareness of S&P

Roepke and Schroeder (2019) also found that over 60% of the games represented in the 183 papers they reviewed were targeted towards end-users with no prior knowledge or skill in computer science. In another broad review of games to improve end-user awareness of S&P, Alotaibi *et al.* (2016) argued that games can be effective at improving

end-user awareness of S&P but that most such games focus on general awareness of S&P threats and defensive strategies; few focus on in-depth awareness raising of specific S&P issues such as, e.g., constructing strong passwords. However, there are some exceptions.

Alqahtani and Kavakli-Thorne (2020) presented a design and preliminary evaluation of a mobile augmented reality game, CybAR. Participants downloaded a mobile gaming application and scanned QR codes placed throughout a university. Each QR code was associated with a game challenge to raise participants' awareness of a particular S&P issue. For each challenge, participants were presented with a scenario and asked what they would like to do in response. If they made the wrong choice, they would be presented with the potential security consequences of this choice (e.g., having their account "hacked"). The authors conducted a preliminary evaluation of the game with 91 volunteers at a university. They found that participant impressions towards the game were generally favorable, but there was no empirical test to assess if the game raised awareness.

To raise awareness of the risks associated with malware and weak passwords, Alotaibi *et al.* (2017) presented the design of two mobile gaming applications, Malware Guardian and Password Protector, respectively. Malware Guardian presented users with new malware threats over time, requiring participants to select the correct way to defend against the threat. Password Protector scored participant passwords to help them create strong, memorable passwords. Their work, however, was also not formally evaluated.

While cybersecurity awareness games are typically computer or video games, as exemplified above, Denning *et al.* (2013) designed a board game, titled "Control-Alt-Hack", to "increase people's prioritization of computer security" and "increase people's awareness of computer security needs and challenges." A peripheral goal was also to improve people's awareness of the job responsibilities of computer security professionals. While Control-Alt-Hack was also not formally evaluated, it has been the foundation for follow-up work exploring games as a vector to improve end-user S&P awareness.

From our own review of the literature, while there have been a number of games developed to improve end-user S&P awareness, few have been empirically evaluated.

4.2 Improving Motivation

Motivational interventions explore strategies to make people *want* to act in line with expert S&P recommendations. Prior work seeking to improve end-users' motivation to adhere to expert-recommended S&P advice fall into three broad categories: nudges, transformational games, and pro-social design.

4.2.1 Nudges / soft-paternalism

In their canonical book, *Nudge*, Sunstein and Thaler define nudges as "any form of choice architecture that alters people's behavior in a predictable way without restricting options or significantly changing their economic incentives" (Thaler and Sunstein, 2009). Applied to the context of end-user security and privacy, then, nudges can be viewed as UI design decisions that subtly but predictably preference users towards adhering to expert S&P recommendations. The exploration of S&P design nudges has been a topic of sustained interest in the usable privacy and security community for some time.

Indeed, Acquisti *et al.* (2017) conducted a systematic literature review of extant work at applying nudges and soft paternalism in the context of end-user privacy and security. They start by describing, from a behavioral economics perspective, the challenges end-users face with respect to privacy and security decision making: i.e., information asymmetry, bounded rationality, and a suite of cognitive biases ranging from loss aversion to hyperbolic time discounting to status quo bias. They then introduce dimensions of nudges that have been and can be used to help counteract these challenges. Specifically, they list six such dimensions in total: *information*, in which users are educated or given feedback about the risks associated with given decisions; *presentation*, in which information is shown to users in a manner that reduces cognitive load to understand risk and safe options; *defaults*, in which controls are set to secure and/or private settings by default; *incentives*, in which users are given rewards to behave in accordance with stated preferences; *reversibility*, in which errors are both made less likely and can be quickly corrected; and, *timing*, in which nudges are delivered at

appropriate times. The authors also briefly discuss the ethics of nudging and enumerate a set of four questions that help identify conflicts among nudgers and nudgees and that cause nudge designers to critically reflect on the appropriateness of a given nudge for a given user population.

Given that an extensive review of security and privacy nudges already exists, our goal in this section will not be to exhaustively categorize all extant S&P nudges. Rather, whereas Acquisti *et al.* (2017)'s typology deconstruct nudges along their atomic design elements, our goal will be to categorize these nudges based on their impact on end-user motivation specifically. To that end, we have identified three distinct types of nudges: nudges that *increase friction* for engaging in insecure behaviors; nudges that *encourage engaging* in secure behavior; and, nudges that inform, alarm, or surprise. Note that these categories are not necessarily mutually exclusive; some nudges may, for example, both inform and add friction to selecting insecure options.

Frictions: Nudges that increase friction for engaging in insecure behaviors

In his book, *The Design of Everyday Things*, Don Norman recommends making unsafe or irreversible actions difficult — but not impossible — for users to do (Norman, 2013). Distler *et al.* (2020) operationalize this idea in the context of S&P with "security-enhancing frictions" — interaction designs that complicate or encumber engaging in insecure behaviors in order "to mitigate the risk of a certain attack." A broad range of motivational nudges might be considered security-enhancing frictions that make insecure behaviors less convenient, thereby favoring more secure options.

For example, owing to habituation, users commonly ignore computer security warnings. Bravo-Lillo *et al.* (2013) designed and tested a number of "inhibitive attractors" to combat habituation effects in the presentation of computer security dialogs to end-users. Inhibitive attractors are user interface elements that command end-user attention and inhibit click-through actions until the attractors are heeded. Examples include animations that highlight critical information in security dialogs, progressive reveals that do not immediately afford users the

ability to click through the dialog, and requiring users to type in critical information into text fields before enabling the click-through option. In an experimental evaluation, the authors found that users who were presented with inhibitive attractors were significantly less likely to select the less secure option than those who were shown an unmodified computer security dialog. In a similar vein, in a field trial with 28 participants, Wang *et al.* (2014) evaluated how introducing a delay to posting content on Facebook might allow users to reconsider posting content that may potentially result in regret. The authors found that while such a nudge did indeed result in a number of post cancellations and edits, some users found the nudge to be too intrusive.

Felt *et al.* (2015) reported on a field study they conducted in redesigning Google Chrome's default SSL warnings to improve end-users' comprehension of the risks associated with expired SSL certificates, as well as to improve end-user's adherence to the warning by exiting out of the suspect webpage. Specifically, they removed technical jargon from the warning description, made the safe option easy-to-access and visually prominent, and increased friction for ignoring the warning by hiding the "unsafe" option within a menu. While their proposed re-designs failed to improve end-user comprehension, they nearly doubled the number of end-users who adhered to the "safe option" of the warning.

On balance, while frictions can encourage compliance with expert recommendations, they should be used sparingly: as might be predicted, evidence suggests that they are effective, but annoying.

Sparks: Nudges that encourage and/or simplify engaging in secure behaviors

Fogg (2009) refers to triggers that enhance one's motivation to act as a "spark." In the context of S&P, researchers have also explored nudges that make secure behaviors easier to do, more compelling, and/or attention capturing, thereby reducing the amount of motivation users may need to accept S&P advice.

Incentives can sometimes improve end-user motivation. Through a controlled experiment with Mechanical Turk workers, Nicholson *et al.* (2018) found that financial incentives can also lead to improved motivation to adhere to expert-recommended advice in password creation.

Specifically, users who were offered a modest financial incentive to create longer passwords were significantly more likely to do so than users who were not.

Deferral actions that require users to commit to future action can also work. Frik *et al.* (2019a) reported on the results of two experiments assessing the effects, on S&P decision making, of using commitment devices, or promises to engage in pro-S&P behaviors at some time in the future, as a means to reduce present-bias, or the tendency to discount future risks and gains in favor of immediate gratification. In the first experiment, participants were presented with a dialog to enable automatic updates. A control group was given only two options, to update immediately or to ignore the warning. Two treatment groups were also given the option to be "reminded of" or to "commit to" enabling automatic updates in the future. Compared to control, both treatments were significantly less likely to ignore the dialog — a 12% reduction for the commitment condition, and a 50% reduction for the reminder condition. In a second experiment, the authors tested the effectiveness of these treatments for three S&P behaviors: automatic updates, enabling two-factor authentication, and enabling automatic backups. Their results suggest that the reminder nudge is effective for all behaviors, but the commitment nudge is only effective for enabling automatic updates and enrolling in two-factor authentication scenarios.

In a similar vein, Story *et al.* (2020) presented results from an experiment in which they evaluated how the use of two nudges could increase adoption of secure mobile payment services like Apple Pay. The first nudge was an informational nudge designed to inform end-users of the risks of credit card fraud and how secure mobile payments can address those threats. The second nudge was an "implementation intention" nudge that helped users formulate a plan of where they might use Apple Pay instead of a credit card. In an experiment with 411 participants, the authors found that, relative to a control group of participants who were shown neither nudge, participants who were shown one or both nudges were significantly more likely to have set-up and used Apple Pay in the week following the experiment.

The messaging in nudges can also increase or decrease user motivation. Golla *et al.* (2021) explored how messaging in warnings can affect

Facebook user's adoption of two-factor authentication. In one experiment, they found that emphasizing account security as a user's personal responsibility resulted in 33% more viewers clicking on a notification to enable two-factor authentication. In a follow-up experiment, they further found that addressing users by their first name increased clicks by 26%, that providing an explicit "not now" option increased clicks by 11%, and that blocking notifications increased click rates by 22%. Almuhimedi *et al.* (2015) designed and evaluated a privacy nudge on top of an Android permission manager that, once a day, informed users about the number of times their location, calendar, call logs, or contacts had been accessed by the apps they had installed on their phones. Through a field study, they found that users were significantly more likely to check their permission managers and update their permission settings when presented with these nudges versus when they only had the permission manager installed. These nudges, however, were not presented specifically when a specific app used a permission but were aggregated and shown to users at a randomly selected time throughout the day.

However, motivating nudges can also potentially backfire by increasing user S&P concern over the nudging service provide. In an observational study of 382 German Facebook users, however, Kroll and Stieglitz (2021) warn that nudges that encourage users to engage in pro-S&P behaviors can enhance end-user privacy concern rather than make them feel at ease. Unfortunately, this finding suggests a potential incentive misalignment between first-party data aggregators and their end-users: data aggregators who act in the best interests of their users by nudging them towards improved privacy settings risk increasing their users' general privacy concerns as compared to a data aggregator who instead hides privacy options from their end-users.

4.2.2 Transformational Games

Outside of nudges, transformational games are an emerging class of S&P motivation interventions. Whereas traditional games are developed to specifically be entertaining while played, transformational games are designed to eliciting lasting change in the player. For example, educational

games are designed to teach players skills that are transferrable to the real world well beyond the scope of the game, and perspective-taking games are designed to inform players about perspectives that they may not experience in their own lives. As discussed in the previous section, many games developed to encourage user acceptance of expert S&P advice are educational in nature and focus on improving awareness, but a small number of games have also focused on improving user motivation.

For example, Chen *et al.* (2020) designed a transformational game, Hacked Time, to improve end-user self-efficacy and attitudes towards adhering to expert-recommended security and privacy behaviors. Hacked Time puts the player in the role of a time-traveling detecting who helps a college student deal with a security breach by going back in time and solving puzzles. This temporal shifting is designed to help players assess the causal impact of security-relevant decisions on negative outcomes downstream. In a randomized, controlled trial with 178 online participants, the authors found that participants who played Hacked Time had improved security attitudes and self-efficacy with using cybersecurity tools relative to a baseline group of participants who were given the equivalent information through a text document.

Whereas Hacked Time focused on motivating end-users, other work has focused on motivating developers. Dabrowski *et al.* (2015) presented an observational self-assessment of how gamifying a computer security class led to increased student satisfaction. Specifically, the authors describe an approach to teaching offensive computer security skills that requires students to pick a "hacker" pseudonym and play the role of a hacker in a narrative storyline in which they are presented with security challenges and encouraged to solve those challenges faster than other students. The authors surveyed 183 students and conducted 130 "survey interviews" to solicit student feedback. While there was no baseline comparison group, the authors present some self-reported evidence that students appreciated the gamified approach to learning offensive computer security skills.

While games to increase S&P motivation are less common than games to increase S&P awareness, the preliminary evidence suggests that transformational games may be an increasingly important vector to improve end-user S&P motivation.

4.2.3 Pro-social design

In line with the recent focus on modeling the effects of social influences on S&P acceptance, an emerging line of research has explored the design of pro-social S&P technologies to encourage the use of S&P technologies. In an early position paper, Das discusses three critical dimensions for pro-social design in S&P technologies (Das, 2016): *observability*, or making adherence to expert S&P recommendations easily visible to others; *cooperation*, or allowing for collectives of users to work towards mutually beneficial S&P outcomes; and, *stewardship*, or allowing for some users to act in benefit of other users. While pro-social design can impact awareness, motivation, and ability, its effect on increasing end-user motivation or reducing the need for end-user motivation in increasing S&P acceptance is perhaps the most well studied in extant intervention work.

Observability

In social psychology, "social proof" refers to people's tendency to look to others for cues on how to behave in uncertain circumstances (Cialdini, 1987). In other words, people usually conform to what others around them do unless they have firm conviction for breaking the norm. In the context of S&P, Das defines "observability" as the ability for end-users to witness and emulate the S&P behaviors and decisions of others (Das, 2016). Most existing S&P tools and interfaces are not specifically designed to be observable — they tend to be hidden in configuration menus and produce no easily visible alteration to the user experience or user. Some S&P tools, however, are implicitly observable: graphical passwords, for example, or two-factor authentication dongles.

In early work presaging more recent developments in social cybersecurity, DiGioia and Dourish (2005) describes the idea of "social navigation" as a means of improving the observability of S&P behaviors. Social navigation can be thought of as vestiges of other users' behaviors that, for example, show users which S&P functions and controls are most commonly used by other users in a system. While not formally implemented and evaluated, the authors discuss the possibility that this "social navigation" may help motivate improved S&P behaviors. While

DiGioia and Dourish conceptualized social navigation in the context of S&P, their contribution was primarily theoretical. Goecks *et al.* (2009) presented a design and evaluation of two systems that operationalize the concept of social navigation in S&P. Acumen employed social navigation to help users with cookie management: the system aggregates user decisions to accept or reject cookies for a given website. Bonfire employed social navigation to help users with firewall configuration: the system compiles users' Firewall configuration settings for specific software and presents that information to other users to assist in their decision making. Through a limited field deployment, the authors found that while social navigation can help simplify end-user S&P, a number of issues arise. The most prominent such issue is herding behavior, where users uncritically accept the community consensus without understanding why. This challenge arises, in part, because social navigation systems can help address the motivation barrier but does not improve user awareness. Though herding behavior is, in some ways, the point of social navigation, uncritically promoting herding behavior can result in user confusion and can potentially be vulnerable to attacks where malicious users collude to shift community consensus (Goecks *et al.*, 2009).

Other work has empirically evaluated the effects of social observability on S&P behaviors. For example, Egelman *et al.* (2013) performed a multi-phased experiment in which they assessed the effects of password meters on strength of password selection for both important and unimportant accounts. In particular, they explored two treatment conditions relative to a baseline control condition with no meter: (i) showing users a "peer password motivator" that compared the strength of a participant's password to other users (i.e., "your password is stronger than N% of your peers"), (ii) showing participants a non-social meter that visualized the bit-strength of participant's password. They found that while both meters resulted in higher strength password selection for important accounts, they did not observe a statistically significant difference between the peer motivator and the bit-strength motivator.

However, since Egelman *et al.* (2013), other researchers have found evidence suggesting that social password meters can significantly outperform non-social meters in improving end-user password selection. For

example, Ohyama and Kanaoka (2015) presented five "social" password meter designs and evaluated them relative to a non-social control in a study with 700 Japanese user. They found that all social treatments resulted in significantly stronger passwords than the control. Likewise, Dupuis and Khan (2018) explored the effects of social influence on password strength among 48 university students, finding that social influence, when paired with explicit instructions for how to improve password strength, resulted in significantly higher password strength than non-social meters with the same instructions. However, without instructions on how to improve password strength, there was no notable difference in password strength among those who were shown the social versus non-social password meters. In other words, social influence does appear to improve end-user motivation but must be paired with materials to improve ability to be effective at imporving behavior.

Other researchers have explored increasing the social observability of S&P behaviors beyond passwords. Das *et al.* (2014b) conducted a field experiment with 50,000 Facebook users in which they showed some users a notification of how many of their friends used optional security features on Facebook and other users a notification that additional optional security features were available for their use without any corresponding social meta-data. Both notifications included a button that would transport users to a settings page in which they could enable the features for themselves (Das *et al.*, 2014b). They found that social notifications were significantly more likely to be clicked, that the effect increased for users with more friends who used optional security features, and that users in the social notification conditions were more likely to have adopted one of the promoted security features both 7 days and 5 months after being exposed to the notification.

However, the notifications studied by Das *et al.* (2014b) are only possible if metadata that captures the social relationships of users can be correlated with use of S&P technologies. Such social metadata may only be available to a few social networking companies. Krsek *et al.* (2022) explored if social suggestions from experts and/or from the public at large could also impact S&P acceptance. In a controlled experiment, they found that participants were significantly more likely to select more restrictive privacy settings on Facebook if suggested to do so from a

small group of experts or the public at large. Thus, the beneficial effects of improving S&P motivation through social recommendations and suggestions could be approximated even without social meta-data of which of one's immediate social connections adhere to expert-recommended S&P advice.

Cooperation

Das (2016) defines cooperation in the context of pro-social S&P design as design that allows collectives of users to act in concert towards mutually beneficial S&P outcomes. Many existing S&P tools and behaviors are designed to improve individual S&P postures with little direct consideration for others, though an argument could be made that improving one's individual S&P posture could, in turn, increase the cost-to-benefit ratio of undirected, mass-scale attacks. Cooperative S&P design, in contrast, makes explicit how one's own S&P behaviors contribute to a broader group's or collective's S&P goals.

A canonical example of cooperative S&P is onion routing, in which decentralized and distributed collections of nodes route network traffic randomly amongst themselves to make the origin of network requests more opaque to those who do not have a full view of the network (Goldschlag *et al.*, 1999). Tor project[2], for example, provides a browser through which users can employ onion routing to browse the web. Onion routing is cooperative because individual users contribute nodes to the onion routing network in order to use the network to increase their own privacy; in turn, each additional node in the network improves the anonymity and privacy provided by the network for all of its users.

Decentralized peer-to-peer networks are, more generally, also cooperative S&P technologies for that same reason. CoinJoin protocols for cryptocurrencies, for example, allow groups of users to obfuscate payments on public ledgers by mixing transactions together (Maxwell, 2013). If, for example, Alice wants to send $25 to Bob, and Carlos wants to send $15 to Eve, but neither Alice nor Carlos wants people to know to whom they are sending the funds, they can choose to instead send these funds through a CoinJoin protocol that will accept the $25 and

[2] https://www.torproject.org/

$15 from Alice and Carlos, respectively, and send out a series of $5 payments to addresses owned by Bob and Eve. While the amounts sent and received will still be public, the amounts specifically sent from one party to another will be obfuscated. Of course, the more pairs that are included in the CoinJoin, the harder it will be for an external observer to reconstruct payments.

Another example is the Inter-Planetary File System (IPFS): a decentralized content distribution network that allows users and developers to unlock the benefits of cloud storage without entrusting their data to a centralized institution (Benet, 2014). Logas *et al.* (2022) introduced the idea of a "decentralized privacy overlay" (DePO) that is built on top of IPFS. The authors created a system, ImageDePO, that allows users to share images on Facebook without uploading those images onto Facebook. The secret image is stored in IPFS, and a content hash that allows users to find this image is steganographically encoded into a cover image uploaded onto Facebook that is then distributed to a user's friend network. Other ImageDePO users in that friend network can retrieve the secret image from IPFS.

Researchers have created other forms of cooperative S&P technologies that do not specifically rely on decentralized networks. One example is Thumprint, a socially cooperative form of authentication for small, social groups (Das *et al.*, 2017). Typical authentication models, such as passwords, are often inappropriate for small, social groups. Consider the case of a shared smart TV among roomates. Either every roommate must agree upon a shared PIN that all use, or each roommate must register their own authentication credentials with the device. In the former case, there is no access control and revocation is hard without invalidating the secret for everyone. The latter, in contrast, is heavy handed and introduces a layer of social friction where roommates must keep secrets from one another. Thumprint aims to offer a middle ground through shared secret knocks, in which the secret remains constant but its expression varies among individuals. In so doing, individuals are uniquely discernible, allowing for finer-grained access control, but the group as a whole shares the secret (Das *et al.*, 2017).

Wu *et al.* (2022a) introduced a system that allows collectives affected by data breaches and other institutional privacy violations to

cooperatively author demands for redress. Collectives work together to triage privacy concerns of particular import, identify counterparties to act in order to address those concerns, and to generate and rank specific proposed demands those counterparties can take to address the triaged concerns.

Other researchers have formalized design spaces for cooperative S&P systems for specific application domains or more broadly. For example, Chouhan *et al.* (2019) employed participatory design methodologies to develop a model for Community Oversight for Privacy and Security ("CO-oPS") in the context of mobile applications. Their model consists of five inter-related concepts: community, transparency, individual participation, awareness, and trust. They argue that designs that support community oversight of S&P behaviors should aim to build on existing social structures and improve a sense of community; should make behaviors of other community behaviors transparent and support shared accountabilities; should influence proactive individual behaviors; should support both passive learning from and active monitoring of others' S&P behaviors; and, should allow incorporate mechanisms to help users assess the credibility of community-solicited advice. Balancing individual effort and privacy against group needs and benefits was also an important consideration.

Moju-Igbene *et al.* (2022) also employed participatory design methodologies to synthesize a design for cooperative social S&P controls among small, social group. Their design space is comprised of four concepts — social transparency, structured of governance, stakes and responsibility, and promoting pro-group S&P behaviors — and four associated challenges — balancing security versus privacy, combating social friction, mitigating social herding, and minimizing coordination costs. The authors argue that in designing cooperative S&P controls for small, social groups, it is of paramount importance to account for idiosyncratic group dynamics (e.g., families will have very different desires and needs than work colleagues), to balance issues of access and power (e.g., allowing those with lower levels of digital literacy to retain agency), and to provide developer tools to facilitate implementation of social controls so that it is easy to support the design and implementation of social S&P controls (Moju-Igbene *et al.*, 2022).

Stewardship

Das defines stewardship in S&P as design that allows for trusted individuals to act in benefit of others S&P (Das, 2016). Prior work has illustrated that while people often do not feel motivated to adhere to expert-recommended S&P suggestions, many feel accountable for the S&P of their friends and loved ones (e.g., Das *et al.*, 2014a; 2018b). Others have documented evidence of individuals taking on the role of S&P stewards for friends and loved ones (Murthy *et al.*, 2021; Kropczynski *et al.*, 2021). Researchers have explored how to tap into these feelings of accountability by developing systems that allow expert-users to act as S&P stewards for other users.

Mendel *et al.* (2021) explored how older adults might solicit social support for mobile S&P decisions, such as allowing or denying app permission requests, getting feedback on password creation, or assessing the legitimacy of a suspicious email link. They found that older adults appreciated the idea of social support systems to assist them with S&P decisions and behaviors, and could use scaffolded social support to potentially improve their own digital literacy. They also uncovered stumbling blocks in the development of such systems — e.g., that younger helpers can grow impatient with older adults who do not quickly learn from the provided support. Mendel *et al.* (2022) also built predictive models to identify what they call "supportable moments" — i.e., contexts in which users might benefit from proactive assistance to solve security & safety pertinent challenges on their mobile devices. They collected a dataset in which they asked 150 crowd workers to rate their willingness to receive social support across a series of vignettes that illustrated different security/safety challenges that they might encounter in using their mobile phones. They found that user anxiety, openness to social support, self-efficacy, and general security awareness were all strongly predictive factors for identifying supportable moments, and that different models are needed for older adults and younger adults.

Akter *et al.* (2022) reported on the design and evaluation of a mobile app that allows for two-way stewardship, or joint oversight over mobile apps installed and the permission requests thereof. They evaluated the app with parent-teen pairs in the U.S. Their results highlighted a host

of practical design implications for systems that promote cooperation and stewardship in S&P. For example, parents and teens tended to focus on different threats, where teens considered the specific privacy implications of permissions while parents more generally focused on the safety implications of particular apps. They also also observed friction arising from the discrepancy between the egalitarian power relationship implied by the app design and the more hierarchical power relationship between parents and teens in practice — teens felt it was not their place, for example, to monitor their parents' app use. These findings resonate with the design space for social S&P controls proposed by Moju-Igbene *et al.* (2022), who argued that accounting for idiosyncratic power balances among groups is a key consideration in designing S&P controls for groups.

Bonneau *et al.* (2009) introduced the idea of "privacy suites" — pre-configured privacy controls on social networking services that users can automatically adopt from friends or trusted experts. The core motivation behind Privacy Suites is to offload the complexity of privacy configurations to a steward; instead of making many small privacy decisions that are often left at system defaults, users can instead make just one decision: that of who, among a marketplace of potential privacy stewards, to emulate. In doing so, privacy suites reduce the motivation required to improve one's S&P posture on social networking services. The authors formalized the notion of a privacy suite, but the idea was never formally implemented or evaluated.

Stewards have also been explored as means of making S&P systems more accessible for blind and visually impaired (BVI) individuals. Prior work has noted that S&P challenges remain a significant accessibility hurdle for BVI people. In a 2017 global study by WebAIM, for example, 1792 surveyed BVIs ranked CAPTCHAs as the second most problematic daily issue they encountered on the web (WebAIM, 2017). Zhang *et al.* (2021) introduced the idea of "assistive transfer systems" as a way for BVIs to offload the S&P-related accessibility hurdles of the modern web to sighted stewards. As an example, the authors presented the WebAlly system that allows BVIs to solicit assistance from sighted stewards when they encounter inaccessible CAPTCHAs. Trusted stewards are given limited remote control over the section of a BVI's screen that

contains the CAPTCHA challenge and solves the challenge on their behalf. In an evaluation with BVI-Steward pairs, the authors found that both BVIs and their stewards appreciated having a system like WebAlly as a "last resort" option when the BVI was not otherwise able to proceed independently. In the context of CAPTCHAs, assisted transfer systems do not directly improve motivation as CAPTCHAs are not user-chosen S&P behaviors but imposed. More generally, however, assisted transfer systems can help reduce the motivational costs of voluntary S&P behaviors that are otherwise inaccessible (Zhang *et al.*, 2021) — for example, assessing whether or not images taken by BVIs contain potentially private or sensitive information prior to sharing online.

4.3 Improving Ability

Prior work seeking to improve end-users' ability to adopt and implement expert-recommended S&P behaviors also falls into two broad categories: facilitating self-custody & direct manipulation and intelligent interfaces & automation. To date, the vast majority of research in human-centered S&P has focused on identifying and mitigating the ability barrier. A full review of this research is out of scope for this manuscript; rather, we focus on a broad overview of key approaches.

4.3.1 Facilitating self-custody and direct manipulation

In a foundational paper on human-centered security, Cranor asserts that some security touchpoints will always require a human-in-the-loop (Cranor, 2008). In these situations, a vast array of prior work in human-centered security and privacy research focuses on facilitating the self-custody of personal data and digital information by, e.g., designing more intuitive, usable, and fast user interfaces that help end-users accomplish desired S&P goals. Below, we provide a broad overview of prior efforts at facilitating direct manipulation.

Passwords

Simplifying the generation, storage, and memorization of secure passwords has long been a focus of usable security research (Herley *et al.*, 2009). One such thread of research involves helping users memorize secrets. Bonneau and Schechter (2014) implemented a system that trains users to remember strong authentication secrets through chunking and spaced repetition. First, a strong secret is automatically generated and split into distinct chunks. On each login attempt, the user is asked to enter chunks sequentially; if they are unable to do so, then the chunk is shown to them to transcribe. Users are prompted to return and enter their secret at increasingly spaced intervals until they are able to reliably replicate the entire secret from memory. In a randomized, controlled experiment, the authors found that the majority of users who used their system were able to fully memorize their secrets. Das *et al.* (2016) implemented mnemonic password trainers to help users retain these secrets. They found that story-based mnemonic training, in which users embedded secret words in a narrative context of their own construction, helped more users retain their secrets even after extended periods of disuse. Blocki *et al.* (2014) explored the use of space repetition mnemonics to enable the recall of *four* person-action-object passwords over a period of up to 158 days. Participants picked a "famous figure" from a drop-down list and then were randomly assigned action-object tuples. They found that the majority of their participants were able to successfully recall their randomly assigned PAO passwords over time, though the entropy of the action-object tuples were smaller than those tested by the previous two studies.

Another thread of research involves helping users generate better passwords. As we previously discussed, while some password meter designs are intended to motivate users to create stronger passwords, others aim to improve users' ability to create better passwords. For example, Yıldırım and Mackie (2019) evaluated the effects of password creation instructions on the strength and memorability of user-generated passwords in a randomized, controlled trial. They found that, relative to users who were instead simply told to follow a specific password policy

(e.g., must be at least 8 characters, have one lowercase, one uppercase, and one special character), users who were instead given prescriptive instructions on how to generate strong passwords created passwords that were stronger and more memorable after a week and a month. Other researchers have explored more dynamic password creation advice. Shay *et al.* (2015) evaluated the effects of real-time password creation feedback and guidance in an online study with 6,435 participants and found that dynamic guidance can help users create stronger passwords, and can improve user sentiment towards the password creation process, but that multi-step password creation processes can result in users generating weaker passwords. Ur *et al.* (2017) evaluated the effects of dynamic data-driven feedback in password meters on the strength of user-generated passwords. They found that this dynamic feedback can, for some password policies, significantly improve the strength of user-generated passwords without reducing memorability.

A third thread of research involves the design of usable password managers — i.e., software that reduces the memory burden of secure password use for users by automatically managing the generation, storage, and entry of secure passwords. Researchers have proposed a number of improvements to password managers in order to make them easier to use.

One usability challenge with password managers is that many require a strong master password that is used to encrypt other passwords generated and stored by the manager. This master password, in turn, must be strong; if it is compromised, all other passwords in the manager will be compromised as well. However, the generation and memorization of a single strong master password can sometimes be prohibitively difficult for many end-users. To address this issue, McCarney *et al.* (2012) designed and evaluated Tapas, a dual-possession password manager meant to allow for encrypted password storage without the need for a user to remember a master password. Tapas eliminates the need for users to remember a master password by instead storing a secret on a second device — i.e., the user's smartphone. In a usability assessment of Tapas versus a password manager with a master password, the authors found that their participants preferred Tapas and that its benefits came at no statistically significant increase in set-up difficulty.

Some users resist password managers because they are concerned that if their devices are lost or stolen, attackers will be able to easily access all of their accounts managed by a password manager. Guan *et al.* (2017) proposed VaultIME to address this concern: instead of auto-filling one's password, VaultIME will auto-correct passwords that are slightly incorrect. In a simulated evaluation, Guan *et al.* (2017) found that the security loss of this auto-correct functionality is muted given a brute-force attacker with 10 tries, but that many password typos would be corrected. However, they did not evaluate their proposed system with real end-users. Stobert and Biddle (2014) introduced Versipass, a password manager that does not store passwords but cued-recall images that make it easier for users to create and recall strong secrets. Versipass presents end-users with an image that is overlayed with a grid; the selection of specific cells in that grid is then used to deterministically generate a random secret according to a pre-defined policy. Versipass does not store the generated password, but the image and the grid selections; this way, even if a user's device is lost and/or stolen, their passwords cannot be recovered from the disk. Versipass was not evaluated against other password managers, but a cognitive walkthrough surfaced a number of potential usability challenges with the software: e.g., that it was unclear, to users, what was their password.

Observing that many of the core usability challenges with password managers arise from touchpoints in the authentication process that require manual user effort (e.g., generating passwords that adhere to specific policies, difficulties in filling out the correct form field), Stobert *et al.* (2020) introduced ByPass — a password manager that aims to automate the password generation and authentication process for users when interacting with websites that adhere to a pre-specified API. Unlike traditional password managers, thus, ByPass directly interfaces with participating websites so that users have fewer direct touchpoints with the account creation and authentication process. In a usability evaluation of ByPass, the authors found that users generally reacted positively to ByPass.

Another commonly discussed barrier to the widespread adoption of password managers is that they do not facilitate changing passwords. To address this shortcoming, Mayer *et al.* (2016) proposed the use

of crowdsourcable, programming-by-demonstration scripts. The core idea is that a user creates a "recording" demonstrating to a browser plugin how to change a password for a given website; these steps are then recorded and replayed for future such changes. Moreover, these recordings can be shared with other users to effectively crowdsource support for password changes across multiple websites. In a small user evaluation of this concept, the authors found that users found it score high on the system usability scale. However, the concept was not fully implemented or evaluated.

Alternative forms of authentication

Beyond facilitating the generation, memorization, and utilization of secure passwords, researchers have also explored alternative methods of authentication altogether. A full review of the myriad authentication systems designed, implemented, and evaluated is out of scope for this paper; in fact, many such surveys already exist (e.g., Bonneau *et al.*, 2012). Here, we provide a brief overview. Generally, we categorize our overview along the three canonical categories of authentication — what-you-know, what-you-have, and what-you-are.

What-you-know authentication is based on secrets: knowledge of a secret serves as evidence of identity. Passwords, for example, are a form of what-you-know authentication. A key advantage of what-you-know authentication is that it generally is cheap to implement and easy to use — no specialized hardware is necessary, a server need only provide an end-point that users who know a secret can authentication against. A key disadvantage is that secure-use of many forms of what-you-know style authentication requires significant human effort to, e.g., generate and remember strong secrets. Researchers have extensively explored alternative forms of what-you-know authentication that take better advantage of human memory systems.

One such thread of research is on graphical passwords. In general, the key value proposition for graphical passwords is to improve the memorability of strong secrets for end-users by taking advantage of human beings' strong visual memory system. A full survey of prior work on graphical passwords is out of scope for this monograph, and many

such surveys already exist. Biddle *et al.* (2012) present a widely cited and broad survey of prior work on graphical password research spanning their first discussion in the academic literature around 1999 up to 2012. In this survey, they present a three-pronged typology of graphical passwords: recall, recognition, and cued-recall. Recall-based graphical passwords must be reproduced from memory and typically requires users to draw a graphical secret on a blank or gridded canvas. Recognition-based graphical passwords instead require users to memorize graphical prompts and recognize them among a set of decoys when attempting to authenticate. Cued-recall systems require users to construct secrets within existing images — in this way, they do not need to remember a free-form secret from scratch but can instead use landmarks within an image to reconstruct and reproduce secrets.

In a similar vein to graphical passwords, authentication systems that take advantage of people's visual-spatial memory have also been explored. Beyond graphical passwords, visual-spatial authentication encodes secrets as objects or paths to be located within a broader graphical context. For example, Alsulaiman and El Saddik first proposed "3D passwords" as a form of multi-factor authentication (Alsulaiman and El Saddik, 2006; 2008). The core idea was that users would interface with a variety of virtual objects in a 3D virtual world, with each object being mapped to an alternative single-factor form of authentication (e.g., a virtual computer would require the user enter an alphanumeric password or to present a keycard). The end-user's full secret, then, would be interacting with the right virtual objects in the right order. George *et al.* (2019) extended the idea of these 3D passwords to immersive virtual reality systems. In GeoPass, Thorpe *et al.* (2013) encoded secrets as real-world locations and paths on a map and found that these secrets were highly memorable. Das *et al.* (2019b) encoded secrets as a traversed path in a dynamically generated 3D game world, and found that compared to the Android 9-dot authentication system, these visual-spatial secrets were more memorable and secure.

Another thread of research to improve what-you-know based authentication is to take advantage of people's episodic memory: i.e., their memory of more mundane, day-to-day lived experiences. Authentication systems that leverage episodic memory are meant to completely

eliminate conscious memorization of secrets on the part of the user and instead forge secrets out of lived experiences that are idiosyncratic to users. Das *et al.* (2013), introduced autobiographical authentication for mobile phones that would ask users questions about "capturable everyday memories" — i.e., day-to-day experiences that could be sensed or recorded by smartphones, such as where a user might have been at a certain hour or who they called on a given day. They found that while users' memory of their day-to-day experiences were imperfect, the imperfections in their memory could also be modeled to create a form of authentication that was both reliable and easy to scale in difficulty. Hang *et al.* (2015) similarly explored the use of dynamic questions about phone usage for fallback authentication. Through a multi-phased user study, they found that these dynamic questions could, with high reliability, authenticate real users but reject adversarial impersonation attempts. Woo *et al.* (2016) introduced the idea of "life-experience passwords" in which users are authenticated based on their ability to answer a sequence of questions about specific memorable experiences such as with whom and when they took a trip. While significantly slower to enter than alphanumeric passwords, episodic memory based authentication systems reduce active memorization effort and may be appropriate in cases where authentication occurs only infrequently.

Finally, researchers have also explored leveraging muscle memory to improve the memorability of authentication secrets. Wobbrock (2009) proposed rhythm-based passwords with Tapsongs. With Tapsongs, users authenticate by tapping a button to a pre-registered rhythm. In a small user study, Wobbrock found that 83% of legitimate users were able to successfully re-authenticate after registering a rhythm-based password, and that between 10 - 20% of adversaries who had exposure to the secret could successfully reconstruct the password. Drawing inspiration from Tapsongs, Das *et al.* (2017) later created Thumprint, socially-inclusive authentication through shared secret knocks. Like Tapsongs, users first pre-register a rhythm-based secret in the form of a knocking pattern to later authenticate against. However, Thumprint allows small, social groups to authenticate into shared resources as individuals with just one shared secret. Through a lab-based usability and security analysis, they found that Thumprint is broadly resilient against shoulder surfing

attacks and can accurately discriminate between up to 10 legitimate users who all know the same secret.

What-you-have authentication is based on possession: possession of something scarce serves as evidence of identity. House keys, for example, are a form of what-you-have authentication. A key advantage of what-you-have authentication is that people no longer need to construct and keep secure their own authentication secrets; instead, one needs a physical authentication key. A key disadvantage of what-you-have authentication is that if one's key is lost or stolen, then one loses access and potentially makes it easier for an attacker to gain access to resources that are meant to be secure. Moreover, replacing a lost key can be expensive, cumbersome, and slow. In general, there have been few specific academic innovations for what-you-have authentication — advances are usually technical in nature, such as replacing idiosyncratic hardware tokens with general key fobs and smartphones.

What-you-are authentication is based on physical personhood: idiosyncratic attributes of your physical body serves as evidence of identity. Fingerprints and facial recognition systems, for example, are a form of what-you-are authentication. A key advantage of what-you-are authentication is that it mostly eliminates reliance on human memory and ability — instead, one is authenticated passively based on physical characteristics or unconscious behavior patterns. Two key disadvantages of what-you-are authentication are revokability and privacy. If one's fingerprint data is compromised, for example, one cannot revoke access to an attacker without also revoking one's access altogether. Moreover, use of what-you-are authentication requires registering one's biometric information with an authority against whom one must authenticate, which can sometimes invoke institutional privacy concerns. Research on biometric authentication is active, but there has been relatively little work on improving end-user ability with respect to use of biometric authentication specifically, and thus out of scope for this review. Nevertheless, there is ample human-centered research on biometric authentication for the interested reader: e.g., on the ethical and social implications of widespread biometric use (Wickins, 2007) and the usability of consumer-facing biometric deployments (Bhagavatula *et al.*, 2015). Instead, the primary approach to addressing the ability barrier for

what-you-are authentication has been to remove the human-in-the-loop through so-called behavioral biometrics, where users are authenticated based on their idiosyncratic attributes of their actions. We will discuss this approach in more detail in the Intelligent interfaces & automation subsection below.

Beyond the three canonical categories of authentication — what-you-know, what-you-have, and what-you-are — researchers have also proposed a "fourth factor" that is meant to address the ability barrier in authentication systems: "who-you-know", or authentication through vouching. Brainard *et al.* (2006) first proposed a vouching-based authentication system as a form of emergency or fallback authentication for cases where legitimate users can not otherwise authenticate: e.g., if they forget their password or lose their authentication token. They define vouching as peer-level authentication in which one authenticated user vouches for, or authenticates, another unauthenticated user. The authors formally propose how such vouching might be implemented and provide a formal security analysis, but they do not evaluate the concept with real users. In particular, they identify social engineering as a significant risk. Schechter *et al.* (2009) implemented and evaluated the "who you know" approach to fallback authentication with their social authentication system. In their system, users appoint "trustees" who can vouch for them if they lose access to their accounts. Trustees are given account recovery codes that they may grant users after hearing their voice or meeting them in person. To regain access, users must get some configurable number of these account recovery codes (e.g., 2 out of 4). They deployed social authentication to a small subset of users on the Microsoft Live email service and found that the vast majority of participants were able to authenticate into their accounts by reaching their trustees, but that users sometimes forgot who they assigned as trustees (Schechter *et al.*, 2009). In a subsequent pair of security analyses, they found that trustees were broadly resilient to phishing attempts except when an attacker was a close acquaintance. Since this initial work, social authentication has been implemented in a variety of commercial products including Facebook — through its Trusted Contact security feature — and Duo two-factor authentication.

Usable access control & permissions

Access control policies help control permissioned access to specific resources, but are notoriously difficult for end-users to specify, interpret and use. There is a rich tradition of prior work on creating better user interfaces and underlying access control models in order to make it easier for users to directly use and specify access control policies for digital resources.

Some of the earlier work along this vein focuses on simplifying access control policy specification for file systems. For example, Reeder *et al.* (2008) proposed the use of an "Expandable Grid" to make it easier for users to directly specify access control policies for a system. In contrast to a "list-of-rules" approach to access control specification in which users create ever-expanding individual rules to define ideal access control policies, Expandable Grids show users one visual summary of the entire access control policy for a system as matrix in which one axis consists of principals (users or user group), a second axis consists of resources (specific files, programs, etc.), and each cell consists of set of effective permissions for each principal-resource pair. In a scenario-based user study with 36 participants, the authors found that Expandable Grids significantly improved the accuracy and speed of access control policy specification relative to a list-of-rules based approach. Reeder *et al.* (2011) implemented and evaluated a *specificity-precedence* access control conflict resolution method to improve the usability of direct manipulation access control interfaces, such as Expandable Grids. Whereas many access control systems commonly follow a deny precedence conflict resolution method — whereby more restrictive rules always take precedence over less restrictive rules when there is a conflict — the authors found, through a user study with 54 participants, that a specificity precedence method, whereby rules that apply to specific individuals override rules applied to the groups to which those individuals belong, was both more accurate and usable.

Others have explored ways to reduce the acute cognitive burden of specifying entire access control policies at once. For example, Mazurek *et al.* (2011) evaluated the idea of "reactive access control" through which users make ad-hoc access decisions on-demand instead of spec-

ifying full access control policies all at once. Through an experience sampling study, they found evidence that reactive access control supported many of their participants' access control needs — e.g., the desire for greater situational control and interactivity with access control decisions. However, the downsides included participants receiving too many requests and the increased latency of access control decisions. Vaniea *et al.* (2012) reported on a controlled experiment in which they measured the importance of embedding access control configuration interfaces as close as possible to a resource that is being shared instead of all-at-once in a separate screen. For a photo gallery interface, they found users were much more likely to notice and correct access control errors when access control interfaces were placed directly under a photo than when the interface was in a sidebar or on a different screen.

While canonically access control has typically followed a "role-based" underlying model in which users are given roles (e.g., administrator) that dictate their access to protected resources, others have proposed tag-based access control systems to more closely match user mental models of how access control should work. For example, Mazurek *et al.* (2014) presented the design and evaluation of Penumbra, a distributed file system that employs tag-based access control to facilitate the creation of access control policies such as share all photos except those tagged as weird, strange or goofy." Penumbra works on distributed file systems with no central resolver, and is designed in part based on prior user studies to better match people's mental models of access control policies. However, it was not directly evaluated with end-users. Klemperer *et al.* (2012) explored the usability of tag-based access control mechanisms through a controlled lab study with 18 participants. They asked users to tag their own photos for strictly organizational purposes or for organizational and access control purposes. They also presented participants with machine-generated access control rules to catalyze discussion about tag-based access control for photo sharing more broadly. They found that users understood and appreciated tag-based access control for photo sharing, especially when the creation of rules could be facilitated through machine-generated rules. Wang *et al.* (2009) proposed "people-tagging" to allow for distributed attributed-based access control specification more usable in collaborative work contexts. In their system, end-users

can tag each other with attributes, and those attributes in turn beget access to resources. If enough of a quorum is reached for a given user-attribute tuple, then the user is granted an attribute and its associated permissions. While the authors implemented and technically evaluated their system, the tagging approach was not formally evaluated with end-users.

Other researchers have explored how to use computing interfaces to simplify access control for physical spaces and contexts. Bauer *et al.* (2007), presented Grey, a smartphone-based access control system by which users can specify and delegate access to resources over which they have authority. Through a 9-month field trial with 19 participants, the authors found that users could more accurately and securely implement their access control policies with Grey than they could with physical key sharing. They found that while users broadly appreciated the use of Grey — 18 of the 19 participants in the study continued using it after the study ended — even a single failure could discourage adoption of novel access control mechanisms over more familiar ones. They also found that by reducing the overhead associated with creating policy changes, users were more likely to update their ideal access control policies as they evolved over time. Finally, they found that unanticipated and serendipitous uses of Grey encouraged adoption.

Still others have explored improving access control over shared digital resources in the home context. Egelman *et al.* (2008a) introduced Family Accounts as a more seamless form of authentication and access control for multi-user devices in home settings, for example. Family Accounts provide one shared "public" account across all members of a household, along with individual profiles that users could switch into as needed for personal or private tasks. They found that participating families significantly preferred to use Family Accounts over individual profiles and/or only one shared profile.

Finally, while the majority of prior usability research on access control focuses on helping users implement and/or specify ideal access control policies for resources they individually or collectively own, some prior research has focused on helping the users subject to those access control policies more easily negotiate access. To that end, Kapadia *et al.* (2004) attempted to make it easier for users to respond to negative

access control decisions by providing feedback about why access was denied and the easiest way to gain access to the desired resource without leaking information to which the user was not meant to be privy.

Simplifying secure email, messaging, and communication

Whitten and Tygar (1999) was a foundational paper that brought to attention the critical need for human-centered design in consumer-facing S&P interfaces. It also systematically identified the difficulties end-users faced when using the PGP platform for sending secure email — much of which pertains specifically to the ability barrier in that secure email interfaces are often confusing and difficult for novices to use. In the decades since this paper, there has been a wealth of prior research on making secure email and messaging more generally more usable.

In a short review article, Ruoti and Seamons (2019) outlined the core security challenges of email, the core usability challenges of secure email, and how to make email interfaces that are both secure and usable. In analyzing prior research on making usable and secure email interfaces, the authors synthesized a set of six key dimensions for secure and usable email: tight integration of security into existing email and communication workflows; inline tutorials to help users take advantage of security features on-demand; streamlined on-boarding to help recipients easily interact with and start using secure emails; understandable design to allow users to make informed decisions and avoid mistakes; and usable key management software to make it easy for users to store, retrieve and discover the cryptographic keys (e.g., their own private key, their contacts public keys) necessary to use secure email. As a case-study, Ruoti and Seamons refer to their own prior work in which they and co-authors presented a redesign of the user interface to the Private Webmail (Pwm) service that was meant to improve usability and reduce user mistakes with secure email (Ruoti *et al.*, 2016). They introduced an artificial delay to encryption to improve end-user confidence in encryption strength, and used the delay to instruct users about who can read encrypted messages. They also made it clear to users when they were sending encrypted email vs. plaintext email, and added inline contextual tutorials to help users understand how to use secure

email and how secure email works. Through a user evaluation with 51 participants, the authors found that their redesigned interface, Pwm 2.0, was rated highly on system usability and significantly reduced the number of mistakes made sending encrypted email messages compared to a baseline control group that used the original Pwm interface.

Other researchers have focused on addressing one of the largest outstanding barriers to widespread use of encrypted email: key management. Lerner *et al.* (2017) presented Confidante, a secure email client that offloads key management to a third-party key management service — Keybase. Through a user study with 15 lawyers and journalists in which the authors qualitatively compared the usability of Confidante against a secure web-browsing extension, Mailvelope, they found that users were able to complete encrypted email tasks faster and with fewer errors using Confidante, and that offloading key management to Keybase simplified the process enough for many users that they found it comparable to using regular email. Bai *et al.* (2017) explored how offloading key management to third-party services affected end-user perceptions and acceptance of secure messaging. They presented 52 users with different encryption models for secure email. These models varied with respect to how keys were managed and distributed — the registration model emulated a third-party messaging service, such as Apple or WhatsApp, centrally storing and distributing keys for end-to-end encryption; in the exchange model, participants generated keypairs and manually exchanged key information with desired message recipients; finally, in the auditing model, participants registered with a third-party messaging service but external auditors check keys to ensure they match. Participants recognized the security benefits of the exchange model, but found the added convenience of automated key management afforded by the registration model to be sufficient for most everyday use-cases.

Beyond secure email, others have looked into making other forms of secure digital communications more usable. One such example is Do *et al.* (2021c)'s Bit Whisperer system, where physically proximate users can securely share messages with another over tabletop surfaces using surface-bound acoustics. The core idea behind Bit Whisperer is that because sound diffuses more quickly over air than when reflected over a physical surface, it is possible to restrict the domain of acoustic

communication for short-range messages to a physically visible, tangible surface. Adversaries who do not have a device on the same physical surface, thus, would be unable to eavesdrop on messages transmitted over the tabletop surface, while other devices on the surface and within a meter of the transmitting device would be able to. In so doing, users can visibly see exactly with whom they are wirelessly communicating.

Simplifying S&P for under-served populations

The vast majority of research on addressing the ability barrier in S&P, to date, has focused on "users"-at-large: i.e., a monolithic conception of an average user with no special circumstances. More recently, there has been a burgeoning interest in addressing the unique S&P challenges faced by specific, under-served user populations. In outlining this vision in the New Security Paradigms Workshop, Wang (2018) calls this thrust of research inclusive security and privacy — i.e., "designing security and privacy mechanisms that are inclusive to people with various characteristics, abilities, needs and values."

Usable S&P systems for people with visual impairments

Perhaps the most mature thread of this research on inclusive security and privacy is on improving S&P interfaces for people with disabilities — and specifically those with visual impairments (PVIs). Researchers at the intersection of usable S&P and accessibility have introduced and evaluated a broad diversity of designs to make S&P easier for PVIs.

In an early study with more than 150 participants, Bigham and Cavender (2009) found that audio CAPTCHAs —- which are meant to be accessible alternatives to visual CAPTCHAs — are more difficult and time-consuming to solve than visual alternatives. To address this discrepancy, the authors introduced a new interface for solving audioCAPTCHAs that allowed PVIs to control the playback of the CAPTCHA directly in the answer box instead of needing to shift focus back and forth between the playback interface and the answer box. The authors found that their interface improvement increased PVI's success with solving audio CAPTCHAs by 59%. Jain *et al.* (2019) observed that existing audio CAPTCHAs were also insecure against

automatic speech recognition (ASR) attacks. To address both the usability and security shortfalls of existing audio CAPTCHAs, the authors proposed reCAPGen: a system that uses ASR for generating secure and usable audioCAPTCHAs while also helping transcribe audio clips on which state-of-the-art ASR fails. Specifically, reCAPGen creates audioCAPTCHAs by splicing and processing audio clips out of radio programs, podcasts, and videos that ASR failed to accurately transcribe. Through a user study with 60 sighted people and 19 people with visual impairments, they found that when people with visual impairments were asked to transcribe just the last two words of a reCAPGen CAPTCHA with 5-7 words, they were able to successfully complete the task 78% of the time with an average response time of 14.5 seconds — both significant improvements over traditional audio CAPTCHAs.

Also noting the usability discrepancy between visual and audio CAPTCHAs, Fanelle *et al.* (2020) introduced four alternative audio CAPTCHA designs. The Math prototype asked users to perform simple addition and subtraction; the Character prototype asked users to count the occurrence of a specific character within a string of alphanumeric characters; the Pauses prototype, which is a variation of existing alphanumeric audio CAPTCHA designs, asked users to transcribe the alphanumeric characters they heard but incorporated longer pauses between characters to minimize screen reader interference; and, the Categories prototype asked users to count the number of sounds, in a series, that belonged to a certain category (e.g., bird chirps, baby cries). The authors evaluated each of these alternative designs against existing audio CAPTCHAs in a randomized, controlled experiment with 67 people PVIs from around the world over the course of three weeks. The authors found that all four designs were significantly more accurate and faster than the control audio CAPTCHA, but each offered different benefits. The Math and Character prototypes were most accurate (89% and 87%, respectively); the Categories prototype was the fastest to complete; the Math prototype provided the highest resilience against both random guessing and automated speech recognition attacks; however, the Pauses prototype was the most preferred by participants.

Beyond CAPTCHAs, other researchers have explored making mobile authentication for accessible for people with visual impairments. Azenkot

et al. (2012) discovered that many people with visual impairments (PVIs) did not use optional authentication methods to protect their smartphones, partially because many mobile authentication systems are disproportionately cumbersome for these users. In response, they introduced and developed PassChords — a non-visual authentication method for touch surfaces. Users enter PassChords by tapping on a touch surface with one or more fingers — the secret in the unique sequence of taps and the distinct fingers used for the taps. Through a study with 16 blind participants, the authors found that PassChords were faster than traditional pin-based authentication methods for PVIs and that the entropy of the user-generated passchords roughly approximates the entropy of a 4-digit PIN.

Barbosa *et al.* (2016) introduced and evaluated a novel password manager, UniPass, designed to be more accessible for people with visual impairments. UniPass is a smartphone based password manager that stores login credentials and transfers these credentials to desktop/laptop devices on-demand to simplify the authentication process. In a comparative lab evaluation with 10 PVIS, the authors found that, compared to existing commercial password managers, PVIs were more likely to successfully complete authentication tasks with UniPass, were faster at completing those tasks, and that seven out of ten participants preferred UniPass.

Usable S&P systems for victims of domestic violence

Havron *et al.* (2019) proposed and evaluated clinical computer security, an approach for helping victims of cyberattacks in the context of intimate partner violence (IPV). The authors worked with the New York City Mayor's Office to End Domestic and Gender-Based Violence to deploy a clinical computer security service for IPV survivors. As a part of their consultations, the authors developed tools to help identify spyware on victims' phones, as well as a series of guides to help clients and consultants manually check important security configurations. For 23 of the 44 consultations reported in the paper, the authors identified key security risks, vulnerabilities, or vectors for abuse and provided victims and referring professionals with information about the vulnerabilities and how to address them.

Usable S&P systems for designed for minority communities
Logas *et al.* (2022) introduced Image DePO: a browser extension that allows users to share secret images over Facebook. The image users actually want to share are uploaded onto inter-plantary file system (IPFS), and can be found only through a content hash. This content hash is then steganographically encoded into a "cover image" that is shared to a user's friends on Facebook. Friends that also have Image DePO installed will be able to retrieve the steganographically encoded content hash and, therefore, see the secret image. Facebook and users without the Image DePO extension, in contrast, will see only the cover image. The authors evaluated their Image DePO prototype with 19 BIPOC and LGBTQ+ participants in order to center experiences of under-served populations who are disproportionately surveilled online; their results suggest that participants were broadly appreciative of the concept and expressed an immediate desire for using it.

In general, while there is a wealth of behavioral research identifying the unique S&P challenges faced by specific user populations that correlate with the ability barrier (e.g., sex workers, journalists, older adults, ethnic minorities, LGBTQ individuals, undocumented migrants, victims of domestic violence), there has been less research addressing these challenges to date.

4.3.2 Intelligent interfaces & automation: Keeping humans-out-of-the loop

While making S&P interfaces easier should always be a key priority for usable security researchers and practitioners, sometimes the best way to address the ability barrier is to remove or reduce the role of the human-in-the-loop. Beyond facilitating self-custody & direct manipulation, therefore, the second way researchers have explored addressing the ability barrier is through the use of intelligent interfaces to automate away S&P decisions from end-users or to simplify S&P decision-making. Automating away S&P decisions from end-users has been touted both by usability and security experts. As Jakob Nielsen, one of the originators of the "heuristic evaluation" method that is commonly used in industry to

assess the usability of consumer-facing user interfaces, states: "[attacks] cannot be thwarted by placing the burden on users to defend themselves at all times. Beleaguered users need protection, and the technology must change to provide this" (Nielsen and Alertbox, 2004).

Intelligent assistants, recommenders, and automation that help users with S&P

One approach of increasing popularity is in the creation of intelligent assistants that can automatically configure and/or recommend S&P settings for users based on context, knowledge of end-user preferences, and through collaborative filtering. In contrast to smart home intelligent assistants and recommender systems that monitor personal data to assist users with general tasks, researchers have also explored the creation of intelligent assistants that help users configure their security & privacy settings.

Early approaches focused on helping users configure privacy settings for location-sharing applications and location-based services. For example, Sadeh *et al.* (2009) introduced a selective location sharing app, PeopleFinder, and used it as a testbed to help users more effectively specify their privacy preferences. They evaluated the effectiveness of predicting a user's ideal policy using a machine learning classifier trained on previous decisions, and found that a machine learning based approach can significantly improve the accuracy of user's specified privacy policies (Sadeh *et al.*, 2009). Similarly, Zhao *et al.* (2014) introduced a collaborative filtering approach to help users configure location privacy preferences for location-based services. Assuming a binary decision — Allow or Deny — the authors proposed to make recommendations to users based on what other similar users have decided for other similar decisions: for example, whether or not to allow sharing one's location in the morning while at a restaurant. The authors compared their approach against a suite of standard machine learning models; they found that both collaborative filtering and standard machine learning models can predict user privacy preferences around 75% of the time, however collaborative filtering can do so with less training data than standard machine learning models.

Researchers have also explored using machine learning to automatically configure mobile application permissions more generally. Liu *et al.* (2016b) proposed and evaluated a collaborative filtering approach for helping users specify mobile application permissions. The authors collected an initial set of application permission configurations from a group of crowdworkers; they then employed a collaborative filtering approach in which the preferences of users who were more demographically similar were weighted more strongly than users who demographically differed. In a preliminary user study, the authors found evidence that participants found value in the permission recommendations — after an initial calibration period, many participants opted to reactivate the collaboratively filtered permission recommendations.

Likewise, Liu *et al.* (2016a) presented the implementation and evaluation of a personalized privacy assistant (PPA) to help users configure smartphone application permissions. The authors proposed a methodology to learn distinct privacy profiles that help predict ideal permission configurations for users. To evaluate their approach, the authors conducted a between-subjects field experiment with 72 participants, where the treatment group of participants received permission recommendations pursuant to their closest profile and the control group received no recommendations. The authors found that participants in the treatment group accepted around 79% of the PPA's recommendations that participants in the treatment condition more quickly converged on their ideal privacy settings. Moreover, the majority of participants who received recommendations found those recommendations helpful. Following the success of the aforementioned PPA for smartphone permission configurations, Das *et al.* (2018a) outlined a broader vision for PPAs for the Internet of Things that informs users about the data collection practices of nearby IoT resources and helps users configure their privacy settings to mitigate unwanted data collection. A core piece of their vision involves the use of machine learning to predict end-user privacy preferences and expectations to reduce the number of decisions end-users must make as they navigate through many different contexts and enter the purview of many different devices.

Most recently, Do *et al.* (2021b) introduced the concept of "smart physical privacy barriers" as a means of using intelligent automation

to protect users from unwanted ambient sensing. As a case study, they present "Smart Webcam Cover" — a webcam cover that automatically closes when the LED indicator associated with webcam use turns off or is suppressed. Through a series of user evaluations, the authors found that users strongly preferred Smart Webcam Cover over a manual alternative, and trusted that it provided more complete coverage relative to manual operation.

4.3.3 Context-aware authentication

Researchers have also explored automating away conscious user effort in authentication. The most well studied such approach is through behavioral biometrics — i.e., authentication users based on idiosyncratic attributes of their actions instead of their person. Monrose and Rubin (1997) presented a canonical proof-of-concept for behavioral biometrics by exploring user authentication through keystroke dynamics: i.e., variations in how users type. On a dataset of 42 distinct keystroke dynamic profiles, their best performing algorithm obtained approximately 90% accuracy. Since their work, researchers have explored other instantiations of behavioral biometrics — for example, by authenticating users based on their mouse movements (Jorgensen and Yu, 2011), gait (Mantyjarvi *et al.*, 2005), and even social interaction (Sultana *et al.*, 2014). One emerging use-case for behavioral biometric approaches is to facilitate continuous authentication, where a user need only actively authenticate (e.g., with a secret) once, and then will remain authenticated based on passively observed behavioral biometrics (Stylios *et al.*, 2016).

Other researchers have explored using contextual factors to estimate the risk associated with a given authentication attempt. Hayashi *et al.* (2013), introduced Context-Aware Scalable Authentication, or CASA. The key idea behind CASA is to modulate the difficulty of an authentication challenge based on a contextual assessment of risk — for example, if one is in a low-risk context such as at one's home in the evening, perhaps only a weak form of authentication is necessary; however, if one is in a high-risk context, such as in a foreign country after a period of extended disuse, perhaps a stronger form of authentication is necessary. Through a series of field studies with end-users, the authors found

that users were broadly receptive to the concept of context-modulated authentication; even participants who had previously used no form of mobile authentication reported that they wanted to use CASA: the security benefits it provided in "high-risk" contexts was a sufficient price for the added security without sacrificing usability in low-risk contexts (Hayashi *et al.*, 2013).

Context-awareness has also been used to improve multi-factor authentication. One key reason users reject the adoption of two-factor authentication is because it requires additional active effort that encumbers the authentication process. Accordingly, researchers have explored automating two-factor authentication through contextual signals. Karapanos *et al.* (2015), presented Sound-Proof, a system that eliminates conscious human effort in two-factor authentication by comparing the audio fingerprints of the authenticating device and the second-factor device. If the audio fingerprints match, then the second-factor device is deemed to be in possession of the authenticating user. In an initial user evaluation, they found that participants strongly preferred Sound Proof as a form of passive two-factor authentication over a more traditional, active alternative (Google 2-step verification).

4.3.4 Concerns with automation

The vast majority of research in addressing the ability barrier, to date, has focused on facilitating direct manipulation of security & privacy settings. However, there has been renewed interest in creating intelligent interfaces that offload the burden of personal privacy and security management from users, largely owing to advances in machine learning modeling methods and the broad availability of personal data on which to train rich models. These approaches have shown promise, but it is worth noting that researchers in the space have also critiqued automation and noted that relying too much on eliminating the human-in-the-loop can have unintended and negative consequences that could reduce end-user acceptance of expert-recommended S&P tools and behaviors.

Edwards, Poole, and Stoll warned of the potential harms of automating away end-users from security decisions as early as 2008 (Edwards *et al.*, 2008). They argued that while removing the human-in-the-loop

in security systems may seem like a panacea for security problems, many socio-technical factors may mitigate against the appropriateness and acceptability of automation in the context of end-user security. The authors assert that automation systems have a limited view of social and environmental context and thus are unlikely to accurately predict user intent, trust, or context of use; in turn, they automation systems cannot be expected to make security decisions in line with user expectations and needs. The challenge is further compounded when taking into account the fluidity and nuance of social and work activity. The authors note, however, that their critiques primarily apply to "rigid" automation systems and note that there may be more room for dynamic automation systems that can be flexible and adapt over time.

Wash *et al.* (2014) illustrated this harm in the context of automated software updates. Through a mixed-methods study, the authors explored how individuals auto-update settings matched their intentions and preconceptions about how their computers updated. Owing in part to automation, the authors found that the majority of their participants misunderstood their computer's update settings and that over half of their participants could not execute their intentions for update management. While participants' computers were likely more secure as a result of their confusion, the authors poignantly conclude: "Removing the user from most of the decisions makes it more difficult for the user to intelligently make the remaining decisions that cannot be fully automated." Thus, even tasks that can successfully be automated can have negative security externalities if not paired with accompanying improvements to end-user awareness and motivation.

In short, as advances in artificial intelligence and machine learning promise a future of intelligent interfaces that can relieve end-users from endless S&P decisions, these intelligent interfaces must be thoughtfully deployed so that users can retain agency over their personal devices and data.

5

Discussion

5.1 Summary of the SPAF

In reviewing the prior literature on the challenges end-users face with accepting expert-recommended S&P behaviors, and then relating these challenges to psychological models of human behavior and technology acceptance, we have identified three fundamental barriers to widespread end-user acceptance of expert-recommended S&P behaviors and tools: awareness, motivation, and ability. Taken together, these barriers make up what we call the "Security & Privacy Acceptance Framework" (SPAF). Efforts to mitigate any of these barriers can be said to increase end-user S&P acceptance; efforts that exacerbate these barriers, however unintentionally, can be said to decrease S&P acceptance.

Awareness is knowledge of threats pertinent to the digital resources one wants to protect, and the behaviors and tools that experts recommend to mitigate those threats. Prior research suggests that end-user awareness in the context of S&P can be impacted by social engagement — e.g., hearing about threats through stories from friends and colleagues (Rader *et al.*, 2012), observing others engage in specific security behavior (Das *et al.*, 2014a); their own mental models of how their devices and accounts work and their perceptions of how threats might manifest in

those models — e.g., perceiving adversaries as "digital graffiti artists" who just want to make mischief versus professional cybercrime units (Wash, 2010); media exposure to pertinent threats — e.g., news coverage of large data breaches (Das *et al.*, 2018b); and, S&P warnings and notifications — e.g., information dialogs that inform users of specific S&P risks (Bravo-Lillo *et al.*, 2013). Addressing the awareness barrier has been a longstanding focus of the usable security and privacy community. Common approaches include the use of awareness training campaigns in which members or employees of specific organizations are required to take short educational courses or quizzes that inform them of pertinent S&P threats; simulated attacks in which IT professionals spoof real attacks by, e.g., sending faux phishing emails to create teachable moments for unaware individuals; and, informative games that are meant to teach game players specific S&P concepts — such as what makes for a strong password, or how to detect phish.

Motivation is desire to act in accordance with expert-recommended S&P suggestions to protect one's digital resources against pertinent threats. Motivation can be impacted by subjective norms — i.e., beliefs about how important others find the threat (Das *et al.*, 2014a) to be and/or how one might be perceived for accepting expert-recommended S&P suggestions (Gaw *et al.*, 2006); perceived relative advantage — i.e., a personal cost-benefit assessment of accepting expert-recommend S&P suggestions (Redmiles *et al.*, 2018); trialability — i.e., how easy it is to try and revert use of the suggested tool and/or behavior (Vaniea *et al.*, 2014); and, compatibility — i.e., how well using the suggested tool and/or behavior aligns with one's own perceived posture towards S&P (Dourish *et al.*, 2004). Prior research has explored three broad approaches towards improving end-user motivation in S&P: nudges and soft paternalism, where expert-suggestions are made slightly easier and/or more obvious than other options; transformational games that are meant to communicate not just information about S&P but to change end-user attitudes towards S&P more generally; and, pro-social design that allows users to easily observe other users S&P behaviors, allows groups of users to act cooperatively for mutual S&P benefit, and that allows experts to act as stewards for non-experts.

Ability is converting a specific S&P intention into effective action. Two factors affect ability: system usability and accessibility. In his book, *The Design of Everyday Things*, Don Norman identified two gulfs that can negatively impact user experience in product design: the gulfs of execution and evaluation (Norman, 2013). The gulf of execution refers to the degree to which the interaction possibilities of a system correspond to the intentions of the user and what that user perceives is possible to do with the system. Prior research in usable security and privacy have outlined gulfs of execution in many user-facing S&P interfaces as early as the late 1990s, starting with Whitten and Tygar (1999). The gulf of evaluation refers to degree to which the system provides representations that can be perceived and interpreted in terms of the expectations and intentions of the user (Norman, 2013). Prior research in usable privacy and security has also documented challenges users face in understanding how their actions affect the S&P of their devices, accounts and resources (Wash, 2010; Wash *et al.*, 2014). The accessibility of S&P interfaces can also cause ability barriers: prior research has shown how many common S&P-related actions disproportionately encumber people with disabilities (Bigham and Cavender, 2009; Wang, 2018). To address the ability barrier, prior research has focused predominantly on two threads of research that mirror a classical debate in human-computer interaction: facilitating direct manipulation versus constructing intelligent interfaces. Prior research on facilitating direct manipulation has focused on making S&P controls and systems faster, more simple, and more seamlessly integrated into typical workflows for end-users. Prior research on constructing intelligent interfaces has explored methods to remove or otherwise minimize the footprint of the "human-in-the-loop" by, e.g., predicting and automatically configuring S&P settings on behalf of end-users.

Through our synthesizing and framing of these barriers, we have been able to identify and categorize the varied thrusts of usable privacy and security research writ large but we note that the barriers of the SPAF are not mutually exclusive; awareness impacts motivation and ability, motivation impacts awareness and ability, ability impacts awareness and motivation.

5.2 Using the SPAF: Gaps and opportunities for future research

The usable privacy and security community has made great strides in both understanding end-user challenges with S&P and in exploring ways to address these challenges through improvements to existing designs or by proposing entirely new designs. Despite these efforts, few people accept expert-recommended S&P advice today and there remains significant room for improvement of end-user acceptance of S&P behaviors and tools. We list and expound upon a few promising directions we have identified in our reading of the prior literature. We note, however, that this list is non-exhaustive and heavily filtered through the biases of the authors.

5.2.1 Integrative approaches that address awareness, motivation, *and* ability

One of our key motivations for developing the SPAF was to help identify why many end-users still largely reject expert-recommended S&P advice despite the wealth of interventions that have been proposed, implemented, and positively evaluated in prior art. In our review of prior art on addressing the awareness, motivation, and ability barriers in S&P, we found no prior work that aimed to address all three of the aforementioned barriers at once. The vast majority of prior art focuses on addressing just one barrier, though many may touch on more than one peripherally (see Table 4.1). For example, informational games and campaigns can improve end-users' awareness of specific threats and mitigation strategies, but do not necessarily help with or convince end-users to implement these mitigation strategies. Nudges may motivate users to take pro-S&P action, but do not necessarily help convert intention into action nor help users understand how that action addresses a specific pertinent threat. Simplifying user interfaces for specifying expressive access control policies can enable users to turn intention into action, but does not specifically motivate them to do so independently or help them understand the most pertinent threats to the resources they would like to protect. We argue, therefore, that there is a ripe opportunity to develop novel techniques and tools that take an integrative approach to address all three barriers at once.

There is some evidence from prior work that addressing multiple barriers at once is more effective than focusing on just one. One example is social password meters, in which end-users are shown their password strength relative to others. Egelman *et al.* (2013) initially found that these social password meters had no effect on improving password strength — but their intervention was primarily meant to be a motivational nudge. Later, Dupuis and Khan (2018) found that social password meters could significantly increase password strength, but only when accompanied with a guide for how to improve one's password strength. In other words, a social password meter that also incorporated instructions to address the ability barrier was more effective than one that consisted only of a motivational nudge. Similarly, in prior work exploring why security awareness campaigns often fail to result in behavior change, Siponen (2000) found that campaigns with only descriptive messaging of threats are unlikely to change behavior. Instead, descriptive messaging must be paired with prescriptive information that helps individuals not only learn of pertinent threats but helps them make specific behavioral changes that protect against those threats. In other words, Siponen (2000) makes the case that addressing the awareness barrier alone is not enough, one must also address the ability barrier. In later work, Bada *et al.* (2019) made a similar argument, suggesting that awareness is just one factor that leads to behavior change. They argued that culturally appropriate messaging that drives behavioral intention (i.e., addressing the motivation barrier) is essential for awareness campaigns to drive behavior change.

In Figure 5.1, we summarize the broad approaches to addressing the awareness, motivation, and ability barriers explored by prior work. This figure may help researchers and practitioners brainstorm ways in which their existing designs and strategies for encouraging pro-S&P behaviors can be supplemented to better address all three barriers in the SPAF.

5.2.2 Measuring and formally modeling end-user acceptance of expert-recommended S&P advice

The SPAF is conceptual; a systematization of prior literature examining the factors that affect end-user acceptance of S&P recommendations.

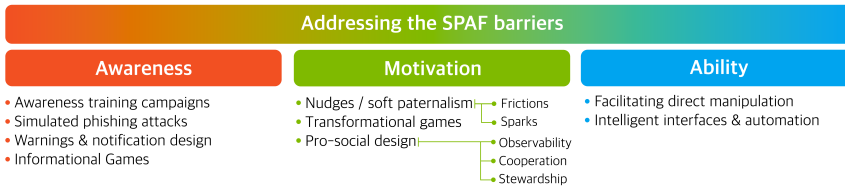

Figure 5.1: Prior work has explored a number of different approaches to addressing one or more of the SPAF barriers. To address the awareness barrier, researchers have explored awareness training campaigns, simulated phishing attacks, warning design, and informational games. To address the motivation barrier, researchers have explored nudges, transformational games, and pro-social design. To address the ability barrier, researchers have explored systems that facilitate direct manipulation and intelligent interfaces.

An intriguing direction for future research, then, is to formalize the SPAF into a predictive model of user acceptance of S&P advice and tools, akin to the Technology Acceptance Model. This would entail, at minimum, a means of measuring S&P acceptance of specific S&P recommendations and open testbeds to study the impact of varied interventions on encouraging and/or discouraging this acceptance.

In terms of measurement, a number of psychometric scales and measures have been developed to assess security behavioral intentions, attitudes and privacy concerns in general. These psychometric scales can be useful in categorizing end-users with respect to S&P. The Westin Index, for example, categorizes people into three broad groups that vary with respect to their general privacy concern: privacy fundamentalists, pragmatists and unconcerned (Kumaraguru and Cranor, 2005). In practice, however, this categorization has been shown to have limited explanatory power in predicting end-user privacy behaviors: in one study, those who were classified as privacy fundamentalists were just as likely those who were classified as unconcerned to share their personal data to receive favorable discounts in an online search engine (Acquisti *et al.*, 2015), for example. The Internet User Information Privacy Concern scale (IUIPC) consists of a series of subscales that help measure how concerned users feel about their personal data being collected, processed and monetized over the Internet (Malhotra *et al.*, 2004). Two other measures are specific to users' concerns about informational privacy

in organizations (Smith *et al.*, 1996) and on mobile devices (Xu *et al.*, 2012). The Privacy Concern Scale (PCS) is a quick read of respondents' general security and privacy concerns (Buchanan *et al.*, 2007). The Security Behavior Intentions Scale (SeBIS) measures user behavioral intention across four areas (Egelman and Peer, 2015a): device securement, updates, password management, and proactive awareness. The six-item Security Attitude scale (SA-6) and its expansion (SA-13) rate participants' general attentiveness to and engagement with security practices (Faklaris *et al.*, 2019; 2022). Longer and more comprehensive measures, although also more burdensome to respondents, are the 31-item Personal Data Attitude measure for adaptive cybersecurity (Addae *et al.*, 2017), and the 63-item Human Aspects of Information Security Questionnaire, or HAIS-Q (Parsons *et al.*, 2017).

While these scales are useful instruments to measure general S&P attitudes and behavioral intentions, there remains an opportunity to develop adaptable measurement instruments that adequately capture a given user's awareness/ability/motivation for specific S&P behaviors and domains, rather than general S&P attitudes, concerns, and behavioral intentions. An instrument that is sensitive enough to capture changes in user awareness/ability/motivation after being exposed to specific interventions may be particularly useful in helping compare the relative efficacy of different proposed interventions.

Beyond measurement, there is a wealth of research on improving end-user security & privacy behaviors, but this research can sometimes be hard to compare: for example, there have been many novel authentication systems proposed, but how can we assess which system users would be most likely to accept and in what contexts? Likewise, there have been many nudges proposed to improve end-user acceptance of S&P advice, but which nudges are most effective and under what conditions? More generally, there is a need for future research to robustly estimate the effects of specific designs and interventions on improving end-user acceptance of S&P advice and tools. An open test-bed that allows for researchers to deploy and assess interventions that are hypothesized to improve S&P acceptance may be particularly fruitful in that regard. Some prior research towards creating such a test-bed exists — e.g., the Security Behavior Observatory at Carnegie Mellon (Forget

et al., 2014); however, to-date, there is no open test-bed that allows for the measurement of end-user S&P acceptance nor the deployment of interventions designed to improve S&P acceptance.

With a measurement instrument that can capture small changes to end-user S&P acceptance after being exposed to specific interventions, and a test-bed that allows researchers and practitioners to assess their proposed interventions in a comparable and controlled environment, one could imagine creating a public repository to benchmark and compare how different designs and interventions might impact end-user acceptance of expert-recommended S&P advice and tools. The combination of this scale, testbed, and public repository should, in turn, accelerate the creation of a science of human-centered security and privacy.

5.2.3 Validating the SPAF and exploring its appropriateness as a stage-model

Our conception of the role of three SPAF barriers in end-user acceptance of expert-recommended security and privacy advice draws from the behavior models discussed in Section 2 and summarized in Table 2.1: i.e., the FBM, the C-HIP for warnings, DoI, and the several permutations of the Technology Acceptance Model. The first two of these models systematize the factors that contribute to people's behavioral response to an interface indicator or message — what Kahneman and others refer to as "fast" thinking, or System 1 of a dual-process cognition model (Kahneman, 2011; Hagger, 2016). An integrative approach to the design and development of interventions, as discussed above, may work best for addressing such non-conscious user behaviors, as the three SPAF concepts will come together in an instant to influence a user's reactions. The second two of these models describe components of more deliberative, rational, "slow" thinking (Kahneman, 2011; Hagger, 2016), such as weighing the value of using a password manager (Pearman *et al.*, 2019), or engaging in the process of software updates (Vaniea and Rashidi, 2016). Models in this second category (also known as "System 2") systematize factors that require more conscious deliberation over time to alter user behavior (e.g., assessing the perceived relative advantage of a system is not something one can instantaneously assess).

The first open question here is to what extent we can reliably measure a user's level of awareness, ability, and motivation for a given expert-recommended security behavior in a model of conscious behavior change. In this manuscript, we have reviewed the prior literature to find empirical evidence of the awareness, motivation, and ability barriers that end-users must overcome to accept and adopt expert-recommended security and privacy behaviors. However, when developing SA-6, Faklaris *et al.* (2019) found that items theorized to measure awareness, motivation, and ability resolved to a single-factor psychometric scale, suggesting that they were too interrelated to be measured independently. Faklaris *et al.* (2019) has since conducted an interview study that identified a common narrative of security adoption segmented into steps, including two steps labeled Threat Awareness and Security Learning that may track with the SPAF. More research is needed to clarify whether awareness, ability, and motivation can be distinctively measured, or whether they are interrelated to the extent that they arise together and co-vary together. This might be accomplished in part by developing more specific measurement tools, as described above. It could also be investigated by borrowing measurement protocols from social-psychology and communication theories that include the SPAF concepts as distinct variables, such as Protection Motivation Theory, Self-Efficacy Theory, or Self-Determination Theory. Answering this question will help point toward whether one SPAF barrier in particular is most important to target for encouraging S&P behavior change.

The second open question is whether the three SPAF concepts are likely to occur in a specific time order. For example, the DoI model describes the Innovation-Decision Process as generally occurring in the time order of (1) Knowledge, (2) Persuasion, (3) Decision, (4) Implementation, and (5) Confirmation. The advantage of knowing the time order is that we can segment users by their step or "stage" of the process, then zoom in and identify the factors that differentiate each segment and that can explain the evolution of thinking and emotions about the target behavior. This avoids a "one size fits all" approach (Egelman and Peer, 2015b) and produces a classification scheme that we can use to design and direct an S&P intervention to those who are most likely to benefit from it. Recent studies of software updates (Vaniea

et al., 2014; Vaniea and Rashidi, 2016) and account sharing (Park *et al.*, 2018; Song *et al.*, 2019; Wang *et al.*, 2022) support the idea that many consciously enacted security and privacy behaviors evolve through time in steps or stages. Faklaris (2022) researched toward identifying a common process model of S&P adoption. If an ideal time order for the SPAF barriers is clarified, we then can identify which interventions are more likely to help people move from one adoption step or stage to the next — and which are *not* likely to help. For example, we can test whether an intervention meant to prompt action (such as a discount code for a third-party password manager) is a better match for a stage associated with motivation vs. one that is associated with awareness.

5.2.4 Bridging the social-technical gap in human-centered security and privacy

Mark Ackerman famously observed that there is often a "divide between what we know we must support socially and what we can support technically" — he termed this divide the "social-technical" gap and argued that addressing this gap was the fundamental intellectual challenge of computer-supported collaborative work (Ackerman, 2000). While Ackerman's paper was addressed to the CSCW and social computing research communities, in particular, his observation holds true for a broad milieu of human-centered security and privacy research as well. Indeed, as a case-in-point, Ackerman highlights the social-technical gap in the Platform for Privacy Preferences Project (P3P) (Cranor, 2003) — an early attempt to create a privacy standard for the web in which users could manage their private information vis-a-vis other counterparties (i.e., other people or institutions). Ackerman argued that static technical privacy controls in P3P fail to adequately capture the social nuance of information sharing decisions.

The presence of this social-technical gap can, in turn, affect all aspects of the SPAF — if technical controls fail to account for social realities, constraints, and preferences, users will be unable to convert intention into action, will be unmotivated to incur immediate social costs for abstract S&P benefits, and will remain unaware of how others navigate similar S&P challenges. Following Ackerman's paper, a number

of researchers in the human-centered privacy and security communities have called for creation of S&P systems and interfaces that are more socially aware, relevant, and intelligent (e.g., Lipford and Zurko, 2012, DiGioia and Dourish, 2005, Das *et al.*, 2014a). Nevertheless, to date, the vast majority of existing S&P interfaces assume individual users making individual decisions about individual accounts and resources. As computing is further enmeshed into the fabric of daily social living — e.g., as sensory devices are increasingly immersed into physical environments, as remote work and collaboration become the norm rather than the exception, and as communication is increasingly facilitated by Internet-connected devices — it is of increasing importance that S&P controls are expressive enough to account for social behaviors and constraints. Indeed, Wu *et al.* (2022b) argued in their systematization of social cybersecurity research: "ignoring human social behaviors in designing S&P systems leads to maladaptive user behaviors that either reduce security, cause social friction, or both. In contrast, by designing for and leveraging human social behaviors in S&P systems, there is an opportunity to both increase the efficacy and the widespread adoption of those systems."

Prior work provides some guidance on how to design S&P tools and controls to address the social-technical gap in human-centered security and privacy. For example, DiGioia and Dourish (2005) called for increased "social navigation" — i.e., the ability for users to observe traces of how other users navigate configuring S&P settings and control to inform their own decisions. Lipford and Zurko (2012) outlined a vision for "community oversight" in which groups of users can collaboratively detect security anomalies and foster greater awareness of and motivation to adhere to expert-recommended S&P advice. Chouhan *et al.* (2019) operationalized what this community oversight might look like in the context of mobile phones through a series of co-design activities. Das (2016) outlined three key design dimensions to design more social cybersecurity systems — observability, or making it easy for bystanders to passively observe when those around them are employing pro-S&P behaviors; cooperation, or allowing groups to act in pursuit of mutually beneficial S&P outcomes; and stewardship, or delegating S&P decisions to trusted trusted stewards who act in benefit of others.

This prior research provides a solid foundation for future work in bridging the social-technical gap in S&P systems, but designing more social cybersecurity and privacy systems will not be a panacea and is likely to introduce its own challenges. For example, in the context of improving individual S&P acceptance, in an observational study exploring how users' use of optional security tools on Facebook correlated with their friends use of those same tools, Das *et al.* (2015) found that when only a few of one's friends use S&P tools like two-factor authentication (2fa), one is less likely to use 2fa than if none of one's friends use it. In other words, social influence could have a negative effect on S&P adoption at low overall levels of adoption. However, the authors also found that S&P systems that were more social by design, such as Facebook's trusted contacts system, were not susceptible to this negative social effect — even at low levels of overall adoption, users were more likely to use trusted contacts if even a few of their friends also used it (Das *et al.*, 2015). In the context of improving group S&P acceptance, Moju-Igbene *et al.* (2022) identified a number of challenges that emerge from or are exacerbated by the design of social cybersecurity systems for small, social groups: for example, power inequalities in governance over shared resources.

In general, while there is a wealth of extant research pointing to the importance of considering social factors in human-centered S&P, there remains a ripe opportunity for future research to develop robust behavioral models of how social factors influence end-user acceptance of S&P, and to draw on those behavioral models to develop novel social cybersecurity systems that improve end-user acceptance of S&P.

5.2.5 Human-AI collaboration to improve end-user S&P acceptance

In discussing prior work seeking to address the ability barrier of the SPAF, we presented a dichotomy that mirrors a classic debate in human-computer interaction research more broadly: direct manipulation vs. intelligent interfaces (Shneiderman and Maes, 1997). Work on direct manipulation aims to make it easier, quicker, or otherwise more amenable for users to manually translate intention into action: e.g., by reducing the cognitive burden of strong authentication (Bonneau and Schechter,

2014) or creating interfaces that make it easier for users to visualize and modify access control policies (Reeder *et al.*, 2008). Work on intelligent interfaces aims to reduce or even eliminate the burden of S&P decision making from users entirely by, e.g., proactively setting privacy controls for users based on learned preferences (Liu *et al.*, 2016a) or automatically authenticating users through behavioral idiosyncrasies that require no active effort (Monrose and Rubin, 1997). The vast majority of research on addressing the ability barrier in human-centered S&P, to date, has focused on improving direct manipulation. However, given the widespread belief that users are the "weakest link" in secure systems (Schneier, 2015), there have been calls to remove the human-in-the-loop from security systems through intelligent automation and stewardship for decades (Nielsen and Alertbox, 2004). Unsurprisingly, then, with recent advances in AI and machine learning, there has been renewed interest in creating novel intelligent interfaces that can automate away many S&P decisions from end-users. Scholars have called into question the appropriateness of intelligent interfaces and automation in the context of human-centered S&P, arguing that automation errors are inevitable and will inevitably lead to user frustration, insecurity, or both (Edwards *et al.*, 2008).

A potentially promising path forward that has been less well explored in the context of human-centered S&P is human-AI collaboration (HAI-C). HAI-C harkens back to J.C.R. Licklider's 1960 treatise on "symbiotic computing" in which human users collaborate and cooperate with intelligent interfaces to improve performance (Licklider, 1960). HAI-C is gaining traction across computing more broadly. Collaborative Human-AI systems, for example, have been used to help radiologists diagnose illness from medical images (Cai *et al.*, 2019), to help customer service representatives respond to an ever-growing volume of support requests (Wang *et al.*, 2012), and to help data scientists run exploratory data analyses (Fast *et al.*, 2018).

In the context of human-centered S&P, rather than striving to entirely remove the human-in-the-loop, a HAI-C approach would instead focus on creating symbiotic AI systems that proactively help people manage S&P decisions while still keeping them informed and aware. One can envision, for example, the creation of proactive intelligent assistants

and agents that understand end-user S&P preferences and goals as they relate to information sharing and work, and proactively help users configure S&P controls for new devices, accounts, and contexts. HAI-C approaches to developing human-centered S&P systems have great potential but have remained relatively under-explored to date. Still, there have been some attempts. One such attempt is Liu *et al.* (2016a)'s personalized privacy assistant (PPA) project that helps users configure their mobile permissions based on their predicted privacy profile. Through a field evaluation, the authors found the the PPA to be effective at helping users configure their mobile app permissions in a manner that was more aligned with their preferences. There remains, however, a ripe opportunity to develop other human-AI symbiotic systems and assistants for human-centered S&P more broadly.

5.2.6 Human-centered threat modeling

Security is a relative term — one is secure, or not, against an adversary with specific capabilities. For example, HTTPS can help prevent one's ISP from being able to parse the information being communicated with Facebook when posting a status update, but does not protect one's content from Facebook itself or the law enforcement bodies that can subpoena Facebook for that information. The adversary that HTTPS protects against, therefore, is third-parties that can intercept or otherwise sniff data packets as they are transmitted over the Internet from sender to receiver. This adversary is often formalized in security work as a "threat model", and almost all security systems have formally codified threat models that they are designed to protect against.

However, to date, threat models are typically formalized based on their technical capabilities and not what end-users are directly concerned about with respect to S&P. This discrepancy contributes to the broad motivation barrier that inhibits widespread end-user acceptance of S&P: S&P are secondary concerns for users, and systems that are designed to protect against abstract threats — however important — are unlikely to inspire users into action. In fact, accepting S&P advice to protect against abstract and seemingly unlikely threats may backfire: prior research has shown that users sometimes believe that use of stringent security

measures are only for paranoid people or people who are performing illicit activities (Gaw *et al.*, 2006). As a result, many users falsely believe that there is no need for them to take protective measures because they "have nothing to hide" (Solove, 2007).

One way to address this discrepancy between what S&P systems protect against and what people actually care about is to incorporate end-user perspectives in threat modeling through one or more of a broad range of HCI methodologies, e.g., scenario-based designs, storyboards assessments, co-design sessions, and semi-structured interviews. While end-users may not have adequate expertise to envision all potential threats, it is imperative that S&P systems address at least one threat that end-users are actually concerned about. If there is no such alignment, there will be a stark motivational barrier to overcome. For example, Logas *et al.* (2022) introduced decentralized privacy overlays (DePOs) to allow individuals to share secret content with their friends on Facebook without uploading that content onto Facebook's servers; while their participants appreciated this functionality, they were more concerned about interpersonal threats (e.g., the wrong friend or family member seeing their post) than they were about Facebook itself. To see widespread organic adoption, therefore, DePOs would have to also help address the more pressing concern on users' minds when sharing content on Facebook: that of expressive audience selection.

5.3 What else matters beyond improving end-user S&P acceptance?

Ultimately, encouraging widespread acceptance of expert-recommended S&P advice is a means to the end of ensuring a secure and trustworthy cyberspace. Thus, we will end with a brief discussion of what prior research suggests are important human factors considerations for future work beyond improving user acceptance of expert-recommended S&P recommendations.

5.3.1 Supporting the wider ecosystem of S&P-relevant actors, decision makers, and stakeholders

The implicit focus of this paper has been on the end-user, but end-users are just one actor in a wider ecosystem of interactors who work to ensure a secure and trustworthy cyberspace. Other pertinent actors include, for example, developers, S&P administrators, and regulators. Prior art suggests the need for future research on supporting these other actors as they make decisions pertinent to ensuring the S&P of the broader computing ecosystem. While a full review of this prior work is out of scope for the current paper, we discuss a few key findings from prior work on each of these stakeholders below.

Developers are responsible for creating the S&P infrastructure and controls underlying the software we all explicitly or implicitly use. In a position paper highlighting the need for more work on making secure development more usable, Green and Smith discuss how mistakes caused by individual developers can have far-reaching S&P consequences (Green and Smith, 2016): "the Heartbleed and Shellshock vulnerabilities led to Internet-wide patch cycles in 2014 and...was caused by an individual developer and affected...millions of users." It is of paramount importance, therefore, to build tooling that helps developers avoid S&P mistakes. Yet, Green and Smith argue, that while usable security research has focused strongly on end-users, comparatively little attention has been paid to improving the usability of secure development workflows (Green and Smith, 2016). They argue that cryptographic APIs and libraries are not developed in a manner that mitigate developer mistakes and assume that all developers have high cryptographic expertise. Other researchers have built on this point, identifying barriers to secure development. Like end-users, for example, S&P are often secondary concerns for developers that compete with more primary concerns in their workflows (e.g., timely product deliveries) and that existing libraries may be both difficult to use and/or assume high-levels of baseline expertise in S&P concepts (Acar *et al.*, 2016). Moreover, many developers are not in full control of their workflows and organizational processes may not adequately support conscious consideration of S&P (Assal and Chiasson, 2019). Social norms and influence can also impact developer use or neglect

of S&P-relevant libraries (Xiao *et al.*, 2014). Following the call for improved developer support, researchers have also begun developing tooling that facilitates the integration of S&P in development workflows (Li *et al.*, 2017; 2021; Jin *et al.*, 2022). However, research in this space remains at early stages and there are many opportunities for further development.

S&P administrators are typically organizational IT staff who identify and respond to attacks on digital infrastructure that could compromise the security and/or privacy of an entire organization — e.g., a university or a corporation. Like developers, administrators have the potential to affect the S&P of broad swathes of users; yet, to date, there has been comparatively little focus on understanding and addressing the challenges faced by administrators. The extant work that has centered administrator experiences suggest that administrators face significant challenges and that existing tools fall short. Haber and Kandogan (2007) reported on results from a 5-year long field study in which they studied the practices of and challenges faced by security administrators across 16 large organizations in the U.S. The authors note that security administrators — more so than other types of system administrators — deal with incredible complexity because they must keep up with a continuously increasing attack surface as new vulnerabilities are discovered, and generally must monitor activity across many different, interacting services and technologies. Despite the importance and complexity of the work done by security administrators, the authors conclude that the tools these administrators use are not designed to support their tasks effectively. Kraemer and Carayon (2007) reported on interviews conducted with 16 network security administrators and security specialists on how human errors factor into the security posture of an organization. The authors showcased the challenges network administrators face in balancing security vs other organizational goals, where exceptions to security policies must often be made to allow for access essential to work goals that would otherwise be considered insecure practice. Tiefenau *et al.* (2020) reported on an interview and subsequent survey study with security administrators to explore common practices and obstacles administrators face with keeping systems under their purview up-to-date. They found that existing tools and infrastructure fail to

properly support even expert administrators, who struggle to assess the consequences of many system updates. For example, updating organizations systems often required significant communication overhead and maintenance overhead as end-users would experience unforeseen complications with their systems after these updates. These barriers, in turn, limited the ability of security administrators to ensure security at scale. Relatively little prior work has explored the creation and evaluation of tools that make S&P administration easier, but one notable exception is work on Let's Encrypt (Aas *et al.*, 2019) and Certbot — an automated certificate authority and command-line configuration tool that makes enabling HTTPS connections on web servers much easier (Tiefenau *et al.*, 2019). The success of Let's Encrypt and Certbot are illustrative of the impact researchers can have on the S&P of the web as a whole by improving the usability of the tooling available to S&P administrators.

Regulators and legislative bodies are tasked with ensuring broad societal compliance with S&P-relevant regulations. In the U.S., for example, the Federal Trade Commission (FTC) has the statutory authority to censure bad actors that act unfairly or deceptively with respect to consumer privacy, but seeks to do so "without unduly burdening legitimate business activity" (Ohlhausen, 2014). While the FTC has emerged as the de facto consumer data protection authority in the U.S. (Hartzog and Solove, 2014), helping set norms for industry data protection practices (Hetcher, 2000; Hans, 2012), it is under-resourced and faces challenges in its ability to keep up with S&P-harms introduced by new technologies (McSweeny, 2018) and in developing regulatory approaches that both allow for beneficial uses of data while meeting the "wide range of consumer preferences for privacy" (Ohlhausen, 2014). The FTC has called for increased research in modeling consumer privacy preferences across contexts and demographics (Ohlhausen, 2014). Moreover, researchers have also begun to explore systems that will help large groups of consumers collectively author privacy grievances that may help the FTC direct their attention (Wu *et al.*, 2022a). Nevertheless, research on connecting affected consumers to the FTC remains sparse, posing a ripe opportunity for future research.

5.3.2 Emerging technologies, emerging threats, emerging opportunities

As emerging technologies — e.g., self-driving cars, virtual reality, deep fakes, bio-hacking, misinformation bots — further enmesh computing into the fabric of society and everyday living, the scope of S&P is steadily expanding in turn. If S&P are fundamentally about protecting users against abuse uses of computing systems, in the near future, S&P problems may no longer be abstract — a security vulnerability could result in physical harm if, e.g., one's self-driving car, implanted sensor, or smart appliance is compromised (e.g., Halperin *et al.*, 2008; Koscher *et al.*, 2010). Likewise, a privacy vulnerability could allow attackers to observe and compromise augmented reality experiences and artifacts that were thought to be personal (Lebeck *et al.*, 2018). Advances in computer-synthesized and generated content — e.g., with deep fakes or misinformation bots — could undermine trust in online discourse and influence public opinion at scale (Chesney and Citron, 2019). Unsurprisingly, then, the standard discourse around the S&P implications of emerging technologies are primarily cautionary. Indeed, as the stakes of insecure computing rise to include material physical harm, it will only become more important to encourage widespread acceptance of expert-recommended S&P. We are already beginning to see many of the aforementioned threats manifest.

However, it is also important to recognize that, with emerging technologies, there will also likely be new opportunities to improve end-user acceptance of S&P by addressing the awareness, motivation, and/or ability barriers. Augmented reality interfaces may be able to more closely pair S&P controls with physical world artifacts, which has been shown to increase user acceptance of the S&P controls that they can leverage (Vaniea *et al.*, 2012). Affective computing devices may be able to improve end-user motivation to respond to pertinent cyberthreats by tapping into corporeal threat sensing mechanisms (Do *et al.*, 2021a; Wilson *et al.*, 2017; Napoli *et al.*, 2020). Intelligent agents may be able to help users implement expressive access control policies without significant manual effort, reducing ability barriers (Liu *et al.*, 2016a). It is unclear which emerging technologies will pose the greatest

opportunity; future research should, however, view these emerging technologies as both S&P challenges and opportunities.

5.3.3 Enhancing consideration of diversity, equity, and inclusion

Much of the research reviewed and summarized in this monograph centers the experiences of privileged individuals: e.g., those in North America or Europe, who are able-bodied, who are wealthy or otherwise fortunate enough to attend top-ranked universities, who feel as though they have "nothing to hide" because they generally fit the expected norms of their cultural or national context. However, a growing body of research examining the S&P needs of underprivileged and marginalized populations is surfacing challenges for a wide spectrum of individuals that are not being adequately addressed by existing expert-recommended systems and advice. Indeed, experts may even be exacerbating these inequities. A full review of this prior work is, again, out of scope for this paper; nevertheless, we cover some key points of interest below as they relate to the SPAF.

Prior research has shown that inequities across demographic and socioeconomic categories can affect awareness of S&P threats and the recommended mitigation measures thereof. For example, through a census-representative telephone survey of 3000 respondents exploring the relationship between demographics and S&P advice sources on the Internet, Redmiles *et al.* (2017) found that less educated users draw from different sources of S&P advice than more educated users, and that users' reported S&P incidents are related to their advice sources. Specifically, users with lower levels of education were slightly more likely to get S&P advice from friends and much less likely to get advice from coworkers or online sources; moreover, those who got advice from coworkers and online sources were less likely to report a negative incident with respect to S&P. In subsequent work, Redmiles (2018) explored inequities in social media privacy behaviors across educational and income strata through a probabilistic telephone survey with 3000 people. The author discovered some evidence to suggest that there may be an "inherited" digital inequality in that higher-income parents and parents who reported higher levels of education were more likely to

help their children set up privacy settings in social media. Individuals in different demographic categories also varied in the types of privacy-preserving behaviors they enacted: e.g., men were more likely than women to turn off cookies, but less likely than women to report using privacy settings on social media.

Prior research also suggests that marginalization and cultural background can correlate with motivation to adhere to expert-recommended S&P advice. One reason for this difference in motivation is the disproportionate harms of S&P violations: minority populations tend to be in more danger of institutional surveillance and the harms that follow. Surveillance scholar Simone Browne, for example, famously observed that there is a disproportionate surveillance of Blackness in the U.S. (Browne, 2015), and others have explored how this manifests in, e.g., predictive policing (Jefferson, 2018). Likewise, scholars have studied the S&P motivations of other minority populations, such as undocumented immigrants and sex workers, and have shown that these populations exhibit a range of behavioral responses — e.g., hyper-vigiliance, resignation, and non-participation — in response to the ever-present threat of institutional surveillance (Guberek *et al.*, 2018; Vannini *et al.*, 2020; McDonald *et al.*, 2021). Similarly, in other regions of the world, journalists and scholars have noted that surveillance tends to focus on specific minority population: e.g., Uighur muslims in China (Beydoun, 2022). Beyond institutional surveillance, some demographic subsets are more prone to interpersonal S&P harms as well: StalkerWare, for example, can harm both men and women but is more commonly used to harm women (Chatterjee *et al.*, 2018). When some populations are more likely to experience S&P harms, or are punished more severely when S&P systems fail, it is unsurprising that motivation to protect oneself from specific harms might vary across demographic contexts.

Cultural influence can also play a role (Dourish and Anderson, 2006): for example, prior research has shown that the most effective messaging in security awareness campaigns varies in different cultural contexts (Bada *et al.*, 2019) — e.g., some cultural contexts respond more to collectivist messaging where improving one's S&P posture is seen more as a social good. Prior work has also suggested that cultural context can affect the frequency of "triggers" that precede S&P behavioral changes,

with users from India being more likely to self-report a social behavioral trigger than users in the U.S. (Das *et al.*, 2019a).

Finally, prior research also suggests that calibrating to and designing for dominant groups can create ability barriers for other groups. One simple example is facial recognition, which is now available as a primary form of authentication for many modern smartphones. Buolamwini and Gebru (2018) illustrated how many commonly used commercial facial recognition algorithms worked less well for Black people and women than for white people and men, in part because those algorithms were predominantly trained on white and male faces. While facial recognition itself raises ethical and privacy questions, this finding suggests that any attempt to use facial-recognition based authentication using these algorithms would inherently privilege white men over other groups. Costanza-Chock (2018) described how an assumption of binary gender norms in safety technology design can result in unintended but significant privacy violations for non-binary and transgender individuals. Finally, while S&P technologies are often inhibiting in nature, the encumbrances they impose are often disproportionately worse for people with disabilities. CAPTCHAs, for example, remain of most significant accessibility hurdles for people with visual impairments (WebAIM, 2017), and even alternatives designed to be more accessible — like audioCAPTCHAs — are often slower and more cognitively demanding for people with visual impairments than visual CAPTCHAs are for others (Bigham and Cavender, 2009; Fanelle *et al.*, 2020).

Recognizing that extant research in S&P often fails to account for the experiences and needs of those on the margins, Wang (2018) outlined a vision for "inclusive security" in which researchers focus on "designing security and privacy mechanisms that are inclusive to people with various characteristics, abilities, needs and values." Indeed, if the ultimate goal is to encourage widespread acceptance of expert-recommended S&P advice, it is of paramount importance that we prioritize research agendas that address the awareness, motivation, and ability barriers for those users who are often neglected but stand to benefit most.

6

Conclusion

How can we encourage end-user acceptance of expert-recommended cybersecurity and privacy behaviors? Addressing this question is a longstanding grand challenge in S&P. Indeed, a 2020 report from McAfee estimated that the global impact of cybercrime exceeded $1 trillion USD annually (Smith and Lostri, 2020). Separately, a 2022 Verizon report on analyzing 5212 security breaches found that the "human element" drove 82% of those breaches (Verizon, 2022). More generally, the S&P behaviors that experts recommend only thinly overlap with the behaviors that people find important and adopt (Ion *et al.*, 2015; Busse *et al.*, 2019). Thus, reducing human error by encouraging or facilitating pro-S&P behaviors remains of broad societal importance.

In this paper, we reviewed prior work in human-centered S&P to systematize barriers to end-user acceptance of pro-S&P behaviors. We found three such barriers: (1) awareness, or the observation that people do not know of relevant security threats and the tools available to protect themselves against those threats; (2) motivation, or the observation that people either do not trust in the efficacy of pro-S&P behaviors to defend against pertinent S&P threats or believe that the costs of implementing those behaviors are too high relative to their benefits;

and, (3) ability, or the observation that people often do not know when, why, and how to accurately implement pro-S&P behaviors, and thus struggle with turning intention into effective action. These three barriers, taken together, make up what we call the "Security & Privacy Acceptance Framework", or the SPAF. We next reviewed the existing body of work in human-centered S&P aimed at increasing security acceptance for the average user. Using the SPAF as a lens, we analyzed why, despite decades of improvements to the usability of end-user S&P systems, widespread acceptance of pro-S&P behaviors remains low: a large majority of prior work to this point has focused on addressing just one the three barriers, typically ability.

Our synthesis of the literature points to a number of fascinating open opportunities and questions for future work. One clear next step is validate the conceptual framework on the SPAF into an empirically supported model with predictive power. Doing so will require creating measurement instruments that can gauge a user's likelihood to accept a given expert-recommended security behavior based on their awareness, motivation, and ability. These scales will need to be situationally customizable, as general attitudes towards security will likely not adequately capture a user's awareness, motivation, and ability for any given S&P behavior. In addition, while many existing interventions in usable S&P have been shown to be effective at addressing one or more of the barriers in the SPAF, there are relatively few interventions that target all barriers. Integrative approaches that target awareness, motivation, and ability together are likely to be more effective at driving end-user acceptance and adoption of pro-S&P behaviors. In Section 5, we further discuss specific opportunities that may be particularly fruitful to explore: i.e., making S&P more social, exploring human-AI collaboration, building a culture of human-centered threat modeling in S&P development, and creating equitable S&P systems that address the SPAF barriers for more than just the "average user."

To conclude, the SPAF helps systematize the barriers that end-users face when considering whether or not to accept or reject expert-recommended S&P advice and behaviors. Moreover, it can be used to help diagnose where extent attempts at improving end-user S&P acceptance falls short and provide insights into how those attempts

can be improved. We urge researchers and practitioners in usable S&P, therefore, to use the SPAF as a lens through which they might diagnose why existing interventions are rejected by end-users, and to explore designs that instead address each of the awareness, motivation, and ability barriers to increase end-user acceptance of pro-S&P behaviors.

Acknowledgments

We would like to thank Youngwook Do for creating the figures we used in this manuscript. We are grateful to our anonymous reviewer who provided helpful comments in improving the clarity of the manuscript. We also thank members and guests of the SPUD, CHIMPS, and the Connected Experiences Lab for their feedback on some of these ideas. Finally, apart from this manuscript relying on the contributions of the thousands of scholars who have contributed towards our collective understanding of why users accept and reject S&P advice, many of the concepts we presented in this work are the cumulative result of conversations we have had with folks in the usable privacy and security community throughout the years.

This work was made possible, in part, by NSF awards CNS-#1704087, SaTC-#2029519, SaTC-#2126058, CAREER-#2144988, the CyLab Security and Privacy Institute, and by the Center for Informed Democracy and Social Cybersecurity.

References

Aas, J., R. Barnes, B. Case, Z. Durumeric, P. Eckersley, A. Flores-López, J. A. Halderman, J. Hoffman-Andrews, J. Kasten, E. Rescorla, *et al.* (2019). "Let's Encrypt: an automated certificate authority to encrypt the entire web". In: *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. 2473–2487.

Acar, Y., S. Fahl, and M. L. Mazurek. (2016). "You are not your developer, either: A research agenda for usable security and privacy research beyond end users". *2016 IEEE Cybersecurity Development (SecDev)*: 3–8.

Ackerman, M. S. (2000). "The intellectual challenge of CSCW: the gap between social requirements and technical feasibility". *Human–Computer Interaction*. 15(2-3): 179–203.

Acquisti, A., I. Adjerid, R. Balebako, L. Brandimarte, L. F. Cranor, S. Komanduri, P. G. Leon, N. Sadeh, F. Schaub, M. Sleeper, *et al.* (2017). "Nudges for privacy and security: Understanding and assisting users' choices online". *ACM Computing Surveys (CSUR)*. 50(3): 1–41.

Acquisti, A., L. Brandimarte, and G. Loewenstein. (2015). "Privacy and human behavior in the age of information". *Science*. 347(6221): 509–514.

Acquisti, A. and J. Grossklags. (2005). "Privacy and rationality in individual decision making". *IEEE security & privacy*. 3(1): 26–33.

Adams, A. and M. A. Sasse. (1999). "Users are not the enemy". *Communications of the ACM (CACM)*. 42(12): 40–46. DOI: 10.1145/322796.322806.

Addae, J. H., M. Brown, X. Sun, D. Towey, and M. Radenkovic. (2017). "Measuring attitude towards personal data for adaptive cybersecurity". *Information & Computer Security*.

Ajzen, I. (1991). "The theory of planned behavior". *Organizational behavior and human decision processes*. 50(2): 179–211.

Akhawe, D. and A. P. Felt. (2013). "Alice in warningland: a large-scale field study of browser security warning effectiveness". In: *Proc. USENIX Sec'13*. 257–272.

Akter, M., A. J. Godfrey, J. Kropczynski, H. R. Lipford, and P. J. Wisniewski. (2022). "From Parental Control to Joint Family Oversight: Can Parents and Teens Manage Mobile Online Safety and Privacy as Equals?" *Proceedings of the ACM on Human-Computer Interaction*. 6(CSCW1): 1–28.

Almuhimedi, H., F. Schaub, N. Sadeh, I. Adjerid, A. Acquisti, J. Gluck, L. F. Cranor, and Y. Agarwal. (2015). "Your location has been shared 5,398 times! A field study on mobile app privacy nudging". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 787–796.

Alotaibi, F., S. Furnell, I. Stengel, and M. Papadaki. (2016). "A review of using gaming technology for cyber-security awareness". *Int. J. Inf. Secur. Res.(IJISR)*. 6(2): 660–666.

Alotaibi, F., S. Furnell, I. Stengel, and M. Papadaki. (2017). "Enhancing cyber security awareness with mobile games". In: *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE. 129–134.

Alqahtani, H. and M. Kavakli-Thorne. (2020). "Design and evaluation of an augmented reality game for cybersecurity awareness (cybar)". *Information*. 11(2): 121.

Alsulaiman, F. A. and A. El Saddik. (2006). "A novel 3D graphical password schema". In: *2006 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems*. IEEE. 125–128.

Alsulaiman, F. A. and A. El Saddik. (2008). "Three-dimensional password for more secure authentication". *IEEE Transactions on Instrumentation and measurement*. 57(9): 1929–1938.

Anderson, B. B., C. B. Kirwan, J. L. Jenkins, D. Eargle, S. Howard, and A. Vance. (2015). "How polymorphic warnings reduce habituation in the brain: Insights from an fMRI study". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2883–2892.

Assal, H. and S. Chiasson. (2019). "'Think secure from the beginning' A Survey with Software Developers". In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.

Auxier, B., L. Rainie, M. Anderson, A. Perrin, M. Kumar, and E. Turner. (2019). "Americans and privacy: Concerned, confused and feeling lack of control over their personal information". *Pew Research Center: Internet, Science & Tech (blog). November*. 15: 2019.

Azenkot, S., K. Rector, R. Ladner, and J. Wobbrock. (2012). "PassChords: secure multi-touch authentication for blind people". In: *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*. 159–166.

Bada, M., A. M. Sasse, and J. R. Nurse. (2019). "Cyber security awareness campaigns: Why do they fail to change behaviour?" *arXiv preprint arXiv:1901.02672*.

Bai, W., D. Kim, M. Namara, Y. Qian, P. G. Kelley, and M. L. Mazurek. (2017). "Balancing security and usability in encrypted email". *IEEE Internet Computing*. 21(3): 30–38.

Barbosa, N. M., J. Hayes, and Y. Wang. (2016). "UniPass: design and evaluation of a smart device-based password manager for visually impaired users". In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 49–60.

Bauer, L., L. F. Cranor, M. K. Reiter, and K. Vaniea. (2007). "Lessons learned from the deployment of a smartphone-based access-control system". In: *Proceedings of the 3rd Symposium on Usable Privacy and Security*. 64–75.

Beautement, A., M. A. Sasse, and M. Wonham. (2008). "The Compliance Budget: Managing Security Behavior in Organisations". In: *Proceedings of the 2008 workshop on New security paradigms - NSPW '08*. New York, New York, USA: ACM Press. 47. DOI: 10.1145/1595676.1595684.

Benet, J. (2014). "Ipfs-content addressed, versioned, p2p file system". *arXiv preprint arXiv:1407.3561*.

Beydoun, K. A. (2022). "The New State of Surveillance: Societies of Subjugation". *Wash. & Lee L. Rev.* 79: 769.

Bhagavatula, R., B. Ur, K. Iacovino, S. M. Kywe, L. F. Cranor, and M. Savvides. (2015). "Biometric authentication on iphone and android: Usability, perceptions, and influences on adoption".

Biddle, R., S. Chiasson, and P. C. Van Oorschot. (2012). "Graphical passwords: Learning from the first twelve years". *ACM Computing Surveys (CSUR)*. 44(4): 1–41.

Bigham, J. P. and A. C. Cavender. (2009). "Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 1829–1838.

Blocki, J., S. Komanduri, L. Cranor, and A. Datta. (2014). "Spaced repetition and mnemonics enable recall of multiple strong passwords". *arXiv preprint arXiv:1410.1490*.

Bonneau, J., J. Anderson, and L. Church. (2009). "Privacy suites: shared privacy for social networks." In: *SOUPS*. Vol. 9. 1–2.

Bonneau, J., C. Herley, P. C. Van Oorschot, and F. Stajano. (2012). "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes". In: *2012 IEEE symposium on security and privacy*. IEEE. 553–567.

Bonneau, J. and S. Schechter. (2014). "Towards reliable storage of 56-bit secrets in human memory". In: *23rd USENIX Security Symposium (USENIX Security 14)*. 607–623.

Brainard, J., A. Juels, R. L. Rivest, M. Szydlo, and M. Yung. (2006). "Fourth-factor authentication: somebody you know". In: *Proceedings of the 13th ACM conference on Computer and communications security*. 168–178.

Bravo-Lillo, C., L. F. Cranor, J. Downs, S. Komanduri, R. W. Reeder, S. Schechter, and M. Sleeper. (2013). "Your Attention Please: Designing security-decision UIs to make genuine risks harder to ignore". In: *Proc. SOUPS'13*.

Browne, S. (2015). *Dark matters: On the surveillance of blackness*. Duke University Press.

Buchanan, T., C. Paine, A. N. Joinson, and U.-D. Reips. (2007). "Development of measures of online privacy concern and protection for use on the Internet". *Journal of the American society for information science and technology*. 58(2): 157–165.

Buolamwini, J. and T. Gebru. (2018). "Gender shades: Intersectional accuracy disparities in commercial gender classification". In: *Conference on fairness, accountability and transparency*. PMLR. 77–91.

Busse, K., J. Schäfer, and M. Smith. (2019). "Replication: no one can hack my mind revisiting a study on expert and non-expert security practices and advice". In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*.

Cai, C. J., S. Winter, D. Steiner, L. Wilcox, and M. Terry. (2019). "Hello AI: uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making". *Proceedings of the ACM on Human-computer Interaction*. 3(CSCW): 1–24.

Carre, J. R., S. R. Curtis, and D. N. Jones. (2018). "Ascribing responsibility for online security and data breaches". *Managerial Auditing Journal*.

Chandrasekaran, V., C. Gao, B. Tang, K. Fawaz, S. Jha, and S. Banerjee. (2020). "Face-off: Adversarial face obfuscation". *arXiv preprint arXiv:2003.08861*.

Chatterjee, R., P. Doerfler, H. Orgad, S. Havron, J. Palmer, D. Freed, K. Levy, N. Dell, D. McCoy, and T. Ristenpart. (2018). "The spyware used in intimate partner violence". In: *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE. 441–458.

Chen, T., M. Stewart, Z. Bai, E. Chen, L. Dabbish, and J. Hammer. (2020). "Hacked Time: Design and Evaluation of a Self-Efficacy Based Cybersecurity Game". In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 1737–1749.

Cherepanova, V., M. Goldblum, H. Foley, S. Duan, J. Dickerson, G. Taylor, and T. Goldstein. (2021). "LowKey: leveraging adversarial attacks to protect social media users from facial recognition". *arXiv preprint arXiv:2101.07922*.

Chesney, B. and D. Citron. (2019). "Deep fakes: A looming challenge for privacy, democracy, and national security". *Calif. L. Rev.* 107: 1753.

Chouhan, C., C. M. LaPerriere, Z. Aljallad, J. Kropczynski, H. Lipford, and P. J. Wisniewski. (2019). "Co-designing for community oversight: Helping people make privacy and security decisions together". *Proceedings of the ACM on Human-Computer Interaction*. 3(CSCW): 1–31.

Cialdini, R. B. (1987). *Influence*. Vol. 3. A. Michel Port Harcourt.

CJ, G., S. Pandit, S. Vaddepalli, H. Tupsamudre, V. Banahatti, and S. Lodha. (2018). "Phishy-a serious game to train enterprise users on phishing awareness". In: *Proceedings of the 2018 annual symposium on computer-human interaction in play companion extended abstracts*. 169–181.

Costanza-Chock, S. (2018). "Design justice: Towards an intersectional feminist framework for design theory and practice". *Proceedings of the Design Research Society*.

Cranor, L. F. (2008). "A framework for reasoning about the human in the loop".

Cranor, L. F. (2003). "P3P: Making privacy policies more useful". *IEEE Security & Privacy*. 1(6): 50–55.

Dabrowski, A., M. Kammerstetter, E. Thamm, E. Weippl, and W. Kastner. (2015). "Leveraging competitive gamification for sustainable fun and profit in security education". In: *2015 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 15)*.

Das, A., M. Degeling, D. Smullen, and N. Sadeh. (2018a). "Personalized privacy assistants for the internet of things: Providing users with notice and choice". *IEEE Pervasive Computing*. 17(3): 35–46.

Das, S. (2016). "Social cybersecurity: Understanding and leveraging social influence to increase security sensitivity". *it-Information Technology*. 58(5): 237–245.

Das, S. (2017). "Social Cybersecurity: Reshaping Security Through An Empirical Understanding of Human Social Behavior". *PhD thesis*. Carnegie Mellon University.

Das, S., L. A. Dabbish, and J. I. Hong. (2019a). "A Typology of Perceived Triggers for End-User Security and Privacy Behaviors". In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 97–115.

Das, S., E. Hayashi, and J. I. Hong. (2013). "Exploring capturable everyday memory for autobiographical authentication". In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 211–220.

Das, S., J. Hong, and S. Schechter. (2016). "Testing Computer-Aided Mnemonics and Feedback for Fast Memorization of High-Value Secrets". In: *2016 Usable Security (USEC) Workshop*.

Das, S., T. H.-J. Kim, L. A. Dabbish, and J. I. Hong. (2014a). "The effect of social influence on security sensitivity". In: *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. 143–157.

Das, S., A. D. Kramer, L. A. Dabbish, and J. I. Hong. (2014b). "Increasing Security Sensitivity With Social Proof: A Large-Scale Experimental Confirmation". In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. New York, New York, USA: ACM Press. 739–749. DOI: 10.1145/2660267.2660271.

Das, S., A. D. Kramer, L. A. Dabbish, and J. I. Hong. (2015). "The role of social influence in security feature adoption". In: *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 1416–1426.

Das, S., G. Laput, C. Harrison, and J. I. Hong. (2017). "Thumprint: Socially-inclusive local group authentication through shared secret knocks". In: *Proceedings of the 2017 chi conference on human factors in computing systems*. 3764–3774.

Das, S., J. Lo, L. Dabbish, and J. I. Hong. (2018b). "Breaking! A Typology of Security and Privacy News and How It's Shared". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. New York, New York, USA: ACM Press. 1–12. DOI: 10.1145/3173574.3173575.

Das, S., D. Lu, T. Lee, J. Lo, and J. I. Hong. (2019b). "The memory palace: Exploring visual-spatial paths for strong, memorable, infrequent authentication". In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 1109–1121.

Davis, F. D. (1989). "Perceived usefulness, perceived ease of use, and user acceptance of information technology". *MIS quarterly*: 319–340.

Denning, T., A. Lerner, A. Shostack, and T. Kohno. (2013). "Control-Alt-Hack: the design and evaluation of a card game for computer security awareness and education". In: *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. 915–928.

Dhamija, R., J. D. Tygar, and M. Hearst. (2006). "Why phishing works". In: *Proc. CHI '06*. No. April. New York, New York, USA: ACM Press. 581–590. DOI: 10.1145/1124772.1124861.

DiGioia, P. and P. Dourish. (2005). "Social navigation as a model for usable security". In: *Proc. SOUPS '05*. New York, New York, USA: ACM Press. 101–108. DOI: 10.1145/1073001.1073011.

Distler, V., G. Lenzini, C. Lallemand, and V. Koenig. (2020). "The framework of security-enhancing friction: How UX can help users behave more securely". In: *New security paradigms workshop 2020*. 45–58.

Do, Y., L. T. Hoang, J. W. Park, G. D. Abowd, and S. Das. (2021a). "Spidey Sense: Designing Wrist-Mounted Affective Haptics for Communicating Cybersecurity Warnings". In: *Designing Interactive Systems Conference 2021*. 125–137.

Do, Y., J. W. Park, Y. Wu, A. Basu, D. Zhang, G. D. Abowd, and S. Das. (2021b). "Smart Webcam Cover: Exploring the Design of an Intelligent Webcam Cover to Improve Usability and Trust". *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 5(4): 1–21.

Do, Y., S. Singh, Z. Li, S. R. Craig, P. J. Welch, C. Shi, T. Starner, G. D. Abowd, and S. Das. (2021c). "Bit Whisperer: Improving Access Control over Ad-hoc, Short-range, Wireless Communications via Surface-bound Acoustics". In: *Proceedings of the 34th ACM User Interface Software and Technology Symposium (UIST)*.

Dodge Jr, R. C., C. Carver, and A. J. Ferguson. (2007). "Phishing for user security awareness". *computers & security*. 26(1): 73–80.

Dourish, P. and K. Anderson. (2006). "Collective information practice: Exploring privacy and security as social and cultural phenomena". *Human-computer interaction*. 21(3): 319–342.

Dourish, P., R. E. Grinter, J. Delgado de la Flor, and M. Joseph. (2004). "Security in the wild: user strategies for managing security as an everyday, practical problem". *Personal and Ubiquitous Computing*. 8(6): 391–401. DOI: 10.1007/s00779-004-0308-5.

Dupuis, M. and F. Khan. (2018). "Effects of peer feedback on password strength". In: *2018 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE. 1–9.

Edwards, W. K., E. S. Poole, and J. Stoll. (2008). "Security automation considered harmful?" In: *Proceedings of the 2007 Workshop on New Security Paradigms*. 33–42.

Egelman, S., A. B. Brush, and K. M. Inkpen. (2008a). "Family accounts: A new paradigm for user accounts within the home environment". In: *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 669–678.

Egelman, S., L. F. Cranor, and J. Hong. (2008b). "You've been warned: an empirical study of the effectiveness of web browser phishing warnings". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1065–1074.

Egelman, S., D. Molnar, N. Christin, A. Acquisti, C. Herley, and S. Krishnamurthi. (2010). "Please Continue to Hold: An empirical study on user tolerance of security delays". In: *Proc. WEIS'10*. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.167.5560.

Egelman, S. and E. Peer. (2015a). "Scaling the security wall: Developing a security behavior intentions scale (sebis)". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2873–2882.

Egelman, S. and E. Peer. (2015b). "The myth of the average user: Improving privacy and security systems through individualization". In: *Proceedings of the 2015 New Security Paradigms Workshop*. 16–28.

Egelman, S., A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. Herley. (2013). "Does my password go up to eleven? The impact of password meters on password selection". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2379–2388.

Faklaris, C. (2022). "Toward a Socio-Cognitive Stage Model of Cybersecurity Behavior Adoption". *PhD thesis*. US National Science Foundation.

Faklaris, C., L. Dabbish, and J. I. Hong. (2022). "Do They Accept or Resist Cybersecurity Measures? Development and Validation of the 13-Item Security Attitude Inventory (SA-13)". *arXiv preprint arXiv:2204.03114*.

Faklaris, C., L. A. Dabbish, and J. I. Hong. (2019). "A Self-Report Measure of End-User Security Attitudes (SA-6". In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 61–77.

Fanelle, V., S. Karimi, A. Shah, B. Subramanian, and S. Das. (2020). "Blind and Human: Exploring More Usable Audio CAPTCHA Designs". In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 111–125.

Fast, E., B. Chen, J. Mendelsohn, J. Bassen, and M. S. Bernstein. (2018). "Iris: A conversational agent for complex tasks". In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

Felt, A. P., A. Ainslie, R. W. Reeder, S. Consolvo, S. Thyagaraja, A. Bettes, H. Harris, and J. Grimes. (2015). "Improving SSL warnings: Comprehension and adherence". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2893–2902.

Felt, A. P., R. W. Reeder, H. Almuhimedi, and S. Consolvo. (2014). "Experimenting at scale with google chrome's SSL warning". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2667–2670.

Ferguson, A. J. (2005). "Fostering e-mail security awareness: The West Point carronade". *Educause Quarterly*. 28(1): 54–57.

Fishbein, M. (1979). "A theory of reasoned action: some applications and implications."

Fishbein, M. and I. Ajzen. (1977). "Belief, attitude, intention, and behavior: An introduction to theory and research". *Philosophy and Rhetoric*. 10(2).

Fogg, B. (2009). "A behavior model for persuasive design". In: *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*. 1. DOI: 10.1145/1541948.1541999.

Forget, A., S. Komanduri, A. Acquisti, N. Christin, L. F. Cranor, and R. Telang. (2014). "Building the security behavior observatory: An infrastructure for long-term monitoring of client machines". In: *Proceedings of the 2014 Symposium and Bootcamp on the Science of Security*. 1–2.

Frik, A., N. Malkin, M. Harbach, E. Peer, and S. Egelman. (2019a). "A promise is a promise: the effect of commitment devices on computer security intentions". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

Frik, A., L. Nurgalieva, J. Bernd, J. Lee, F. Schaub, and S. Egelman. (2019b). "Privacy and security threat models and mitigation strategies of older adults". In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 21–40.

Furnell, S., A. Jusoh, and D. Katsabas. (2006). "The challenges of undersatnding and using security: A survey of end-users". *Computers & Security*. 25(1): 27–35.

Gaw, S., E. W. Felten, and P. Fernandez-Kelly. (2006). "Secrecy, Flagging, and Paranoia: Adoption Criteria in Encrypted E-Mail". In: *Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI '06)*. New York, New York, USA: ACM Press. 591–600. DOI: 10.1145/1124772.1124862.

George, C., M. Khamis, D. Buschek, and H. Hussmann. (2019). "Investigating the third dimension for authentication in immersive virtual reality and in the real world". In: *2019 ieee conference on virtual reality and 3d user interfaces (vr)*. IEEE. 277–285.

Goecks, J., W. K. Edwards, and E. D. Mynatt. (2009). "Challenges in supporting end-user privacy and security management with social navigation". In: *Proceedings of the 5th Symposium on Usable Privacy and Security*. 1–12.

Goldschlag, D., M. Reed, and P. Syverson. (1999). “Onion routing”. *Communications of the ACM*. 42(2): 39–41.

Golla, M., G. Ho, M. Lohmus, M. Pulluri, and E. M. Redmiles. (2021). “Driving 2FA Adoption at Scale: Optimizing Two-Factor Authentication Notification Design Patterns”. In: *30th USENIX Security Symposium (USENIX Security 21)*. 109–126.

Green, M. and M. Smith. (2016). “Developers are not the enemy!: The need for usable security apis”. *IEEE Security & Privacy*. 14(5): 40–46.

Guan, L., S. Farhang, Y. Pu, P. Guo, J. Grossklags, and P. Liu. (2017). “VaultIME: Regaining User Control for Password Managers through Auto-correction”. In: *International Conference on Security and Privacy in Communication Systems*. Springer. 673–686.

Guberek, T., A. McDonald, S. Simioni, A. H. Mhaidli, K. Toyama, and F. Schaub. (2018). “Keeping a low profile? Technology, risk and privacy among undocumented immigrants”. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–15.

Haber, E. and E. Kandogan. (2007). “Security administrators: A breed apart”. *SOUPS USM*: 3–6.

Hagger, M. S. (2016). “Non-conscious processes and dual-process theories in health psychology”. *Health Psychology Review*. 10(4): 375–380.

Halperin, D., T. S. Heydt-Benjamin, K. Fu, T. Kohno, and W. H. Maisel. (2008). “Security and privacy for implantable medical devices”. *IEEE pervasive computing*. 7(1): 30–39.

Haney, J., Y. Acar, and S. Furman. (2021). “"It’s the Company, the Government, You and I": User Perceptions of Responsibility for Smart Home Privacy and Security”. In: *30th USENIX Security Symposium (USENIX Security 21)*. 411–428.

Hang, A., A. De Luca, and H. Hussmann. (2015). “I know what you did last week! do you? dynamic security questions for fallback authentication on smartphones”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1383–1392.

Hans, G. S. (2012). “Privacy policies, terms of service, and FTC enforcement: Broadening unfairness regulation for a new era”. *Mich. Telecomm. & Tech. L. Rev.* 19: 163.

Hargittai, E. and K. Dobransky. (2017). "Old dogs, new clicks: Digital inequality in skills and uses among older adults." *Canadian Journal of Communication*. 42(2).

Hartzog, W. and D. J. Solove. (2014). "The scope and potential of FTC data protection". *Geo. Wash. L. Rev.* 83: 2230.

Havron, S., D. Freed, R. Chatterjee, D. McCoy, N. Dell, and T. Ristenpart. (2019). "Clinical computer security for victims of intimate partner violence". In: *28th USENIX Security Symposium (USENIX Security 19)*. 105–122.

Hayashi, E., S. Das, S. Amini, J. Hong, and I. Oakley. (2013). "Casa: context-aware scalable authentication". In: *Proceedings of the Ninth Symposium on Usable Privacy and Security*. 1–10.

Hendrix, M., A. Al-Sherbaz, and B. Victoria. (2016). "Game based cyber security training: are serious games suitable for cyber security training?" *International Journal of Serious Games*. 3(1).

Herley, C. (2009). "So long, and no thanks for the externalities". In: *Proc. NSPW '09*. New York, New York, USA: ACM Press. 133–144. DOI: 10.1145/1719030.1719050.

Herley, C. (2016). "Unfalsifiability of security claims". *Proceedings of the National Academy of Sciences*. 113(23): 6415–6420.

Herley, C. 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)*. DOI: 10.1007/978-3-642-03549-4_14.

Herley, C., P. C. Van Oorschot, and A. S. Patrick. (2009). "Passwords: If we're so smart, why are we still using them?" In: *International Conference on Financial Cryptography and Data Security*. Springer. 230–237.

Hetcher, S. (2000). "FTC as Internet privacy norm entrepreneur, The". *Vand. L. Rev.* 53: 2041.

Hill Jr, W. A., M. Fanuel, X. Yuan, J. Zhang, and S. Sajad. (2020). "A survey of serious games for cybersecurity education and training".

Ion, I., R. Reeder, and S. Consolvo. (2015). ""...no one can hack my mind": Comparing Expert and Non-Expert Security Practices". In: *Symposium on Usable Privacy and Security (SOUPS)*. 327–346. DOI: 10.1080/0888431022000070458.

Jagatic, T. N., N. A. Johnson, M. Jakobsson, and F. Menczer. (2007). "Social phishing". *Communications of the ACM*. 50(10): 94–100.

Jain, M., R. Tripathi, I. Bhansali, and P. Kumar. (2019). "Automatic generation and evaluation of usable and secure audio ReCAPTCHA". In: *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 355–366.

Jansson, K. and R. von Solms. (2013). "Phishing for phishing awareness". *Behaviour & information technology*. 32(6): 584–593.

Jefferson, B. J. (2018). "Predictable policing: Predictive crime mapping and geographies of policing and race". *Annals of the American Association of Geographers*. 108(1): 1–16.

Jin, H., G. Liu, D. Hwang, S. Kumar, Y. Agarwal, and J. Hong. (2022). "Peekaboo: A Hub-Based Approach to Enable Transparency in Data Processing within Smart Homes". In: *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE Computer Society. 1571–1571.

Jorgensen, Z. and T. Yu. (2011). "On mouse dynamics as a behavioral biometric for authentication". In: *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*. 476–482.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kajzer, M., J. D'Arcy, C. R. Crowell, A. Striegel, and D. Van Bruggen. (2014). "An exploratory investigation of message-person congruence in information security awareness campaigns". *Computers & security*. 43: 64–76.

Kang, R., L. Dabbish, N. Fruchter, and S. Kiesler. (2015). ""My data just goes everywhere": User mental models of the internet and implications for privacy and security". In: *Symposium on Usable Privacy and Security (SOUPS) 2015*. 39–52.

Kapadia, A., G. Sampemane, and R. H. Campbell. (2004). "KNOW why your access was denied: Regulating feedback for usable security". In: *Proceedings of the 11th ACM conference on Computer and Communications Security*. 52–61.

Kaplan, S. A., D. L. Vogel, D. A. Gentile, and N. G. Wade. (2012). "Increasing positive perceptions of counseling: The importance of repeated exposures". *The Counseling Psychologist*. 40(3): 409–442.

Karapanos, N., C. Marforio, C. Soriente, and S. Capkun. (2015). "Sound-Proof: Usable Two-Factor Authentication Based on Ambient Sound". In: *24th USENIX security symposium (USENIX security 15)*. 483–498.

Kelley, P. G., J. Bresee, L. F. Cranor, and R. W. Reeder. (2009). "A "nutrition label" for privacy". In: *Proceedings of the 5th Symposium on Usable Privacy and Security*. 1–12.

Kerner, S. M. (2022). "Colonial Pipeline hack explained: Everything you need to know". URL: https://www.techtarget.com/whatis/feature/Colonial-Pipeline-hack-explained-Everything-you-need-to-know.

Khan, B., K. S. Alghathbar, and M. K. Khan. (2011). "Information security awareness campaign: An alternate approach". In: *International Conference on Information Security and Assurance*. Springer. 1–10.

Klemperer, P., Y. Liang, M. Mazurek, M. Sleeper, B. Ur, L. Bauer, L. F. Cranor, N. Gupta, and M. Reiter. (2012). "Tag, you can see it! Using tags for access control in photo sharing". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 377–386.

Koscher, K., A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, *et al.* (2010). "Experimental security analysis of a modern automobile". In: *2010 IEEE symposium on security and privacy*. IEEE. 447–462.

Kraemer, S. and P. Carayon. (2007). "Human errors and violations in computer and information security: The viewpoint of network administrators and security specialists". *Applied ergonomics*. 38(2): 143–154.

Kroll, T. and S. Stieglitz. (2021). "Digital nudging and privacy: improving decisions about self-disclosure in social networks". *Behaviour & Information Technology*. 40(1): 1–19.

Kropczynski, J., R. Ghaiumy Anaraky, M. Akter, A. J. Godfrey, H. Lipford, and P. J. Wisniewski. (2021). "Examining Collaborative Support for Privacy and Security in the Broader Context of Tech Caregiving". *Proceedings of the ACM on Human-Computer Interaction*. 5(CSCW2): 1–23.

Krsek, I., K. Wenzel, S. Das, J. I. Hong, and L. Dabbish. (2022). "To Self-Persuade or be Persuaded: Examining Interventions for Users' Privacy Setting Selection". In: *CHI Conference on Human Factors in Computing Systems*. 1–17.

Kumaraguru, P. and L. F. Cranor. (2005). *Privacy indexes: a survey of Westin's studies*. Carnegie Mellon University, School of Computer Science, Institute for ...

Kumaraguru, P., J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham. (2009). "School of phish: a real-world evaluation of anti-phishing training". In: *Proceedings of the 5th Symposium on Usable Privacy and Security*. 1–12.

Kumaraguru, P., S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong. (2008). "Lessons from a real world evaluation of anti-phishing training". In: *2008 eCrime Researchers Summit*. IEEE. 1–12.

Landau, S. (2013). "Making sense from Snowden: What's significant in the NSA surveillance revelations". *IEEE Security & Privacy*. 11(4): 54–63.

Lebeck, K., K. Ruth, T. Kohno, and F. Roesner. (2018). "Towards security and privacy for multi-user augmented reality: Foundations with end users". In: *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE. 392–408.

Lebek, B., J. Uffen, M. H. Breitner, M. Neumann, and B. Hohler. (2013). "Employees' information security awareness and behavior: A literature review". In: *2013 46th Hawaii International Conference on System Sciences*. IEEE. 2978–2987.

Lerner, A., E. Zeng, and F. Roesner. (2017). "Confidante: Usable encrypted email: A case study with lawyers and journalists". In: *2017 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE. 385–400.

Li, T., E. B. Neundorfer, Y. Agarwal, and J. I. Hong. (2021). "Honeysuckle: Annotation-guided code generation of in-app privacy notices". *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 5(3): 1–27.

Li, Y., F. Chen, T. J.-J. Li, Y. Guo, G. Huang, M. Fredrikson, Y. Agarwal, and J. I. Hong. (2017). "Privacystreams: Enabling transparency in personal data processing for mobile apps". *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies.* 1(3): 1–26.

Licklider, J. C. (1960). "Man-computer symbiosis". *IRE transactions on human factors in electronics.* (1): 4–11.

Lipford, H. R. and M. E. Zurko. (2012). "Someone to watch over me". In: *Proceedings of the 2012 New Security Paradigms Workshop.* 67–76.

Liu, B., M. S. Andersen, F. Schaub, H. Almuhimedi, S. A. Zhang, N. Sadeh, Y. Agarwal, and A. Acquisti. (2016a). "Follow my recommendations: A personalized privacy assistant for mobile app permissions". In: *Twelfth symposium on usable privacy and security (SOUPS 2016).* 27–41.

Liu, R., J. Cao, K. Zhang, W. Gao, J. Liang, and L. Yang. (2016b). "When privacy meets usability: Unobtrusive privacy permission recommendation system for mobile apps based on crowdsourcing". *IEEE Transactions on Services Computing.* 11(5): 864–878.

Logas, J., A. Schlesinger, Z. Li, and S. Das. (2022). "Image DePO: Towards Gradual Decentralization of Online Social Networks using Decentralized Privacy Overlays". *Proceedings of the ACM on Human-Computer Interaction.* 6(CSCW1): 1–28.

Ma, Y., J. Feng, L. Kumin, and J. Lazar. (2013). "Investigating user behavior for authentication methods: A comparison between individuals with Down syndrome and neurotypical users". *ACM Transactions on Accessible Computing (TACCESS).* 4(4): 1–27.

Malhotra, N. K., S. S. Kim, and J. Agarwal. (2004). "Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model". *Information systems research.* 15(4): 336–355.

Mantyjarvi, J., M. Lindholm, E. Vildjiounaite, S.-M. Makela, and H. Ailisto. (2005). "Identifying users of portable devices from gait pattern with accelerometers". In: *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* Vol. 2. IEEE. ii–973.

Marne, S. T., M. N. Al-Ameen, and M. K. Wright. (2017). "Learning System-assigned Passwords: A Preliminary Study on the People with Learning Disabilities." In: *SOUPS*.

Maxwell, G. (2013). "CoinJoin: Bitcoin privacy for the real world". In: *Post on Bitcoin forum*. Vol. 3. 110.

Mayer, P., H. Berket, and M. Volkamer. (2016). "Enabling automatic password change in password managers through crowdsourcing". *Proc. PASSWORDS. Springer*.

Mazurek, M. L., P. F. Klemperer, R. Shay, H. Takabi, L. Bauer, and L. F. Cranor. (2011). "Exploring reactive access control". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2085–2094.

Mazurek, M. L., Y. Liang, W. Melicher, M. Sleeper, L. Bauer, G. R. Ganger, N. Gupta, and M. K. Reiter. (2014). "Toward strong, usable access control for shared distributed data". In: *12th USENIX Conference on File and Storage Technologies (FAST 14)*. 89–103.

McCarney, D., D. Barrera, J. Clark, S. Chiasson, and P. C. Van Oorschot. (2012). "Tapas: design, implementation, and usability evaluation of a password manager". In: *Proceedings of the 28th Annual Computer Security Applications Conference*. 89–98.

McDonald, A., C. Barwulor, M. L. Mazurek, F. Schaub, and E. M. Redmiles. (2021). "" It's stressful having all these phones": Investigating Sex Workers' Safety Goals, Risks, and Practices Online". In: *30th USENIX Security Symposium (USENIX Security 21)*.

McSweeny, T. (2018). "Psychographics, Predictive Analytics, Artificial Intelligence, & Bots: Is the FTC Keeping Pace?" *Geo. L. Tech. Rev.* 2: 514–514.

Mendel, T., D. Gao, D. Lo, and E. Toch. (2021). "An Exploratory Study of Social Support Systems to Help Older Adults in Managing Mobile Safety". In: *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*. 1–13.

Mendel, T., R. Schuster, E. Tromer, and E. Toch. (2022). "Toward Proactive Support for Older Adults: Predicting the Right Moment for Providing Mobile Safety Help". *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 6(1): 1–25.

Micallef, N., M. Just, L. Baillie, and M. Alharby. (2017). "Stop annoying me! an empirical investigation of the usability of app privacy notifications". In: *Proceedings of the 29th Australian Conference on Computer-Human Interaction*. 371–375.

Milka, G. (2018). "The Anatomy of Account Take-Over". In: *USENIX ENIGMA*. URL: https://www.usenix.org/sites/default/files/conference/protected-files/enigma18_milka.pdf.

Moju-Igbene, E., H. Abdi, A. Lu, and S. Das. (2022). ""how do you not lose friends?": Synthesizing a design space of social controls for securing shared digital resources via participatory design jams,"" in: *Proceedings of the 31st USENIX Security Symposium (SEC)*.

Monrose, F. and A. Rubin. (1997). "Authentication via keystroke dynamics". In: *Proceedings of the 4th ACM Conference on Computer and Communications Security*. 48–56.

Moore, H. and D. Roberts. (2013). "AP Twitter hack causes panic on Wall Street and sends Dow plunging". URL: http://www.theguardian.com/business/2013/apr/23/ap-tweet-hack-wall-street-freefall.

Mostafa, M. and O. S. Faragallah. (2019). "Development of serious games for teaching information security courses". *IEEE Access*. 7: 169293–169305.

Murthy, S., K. S. Bhat, S. Das, and N. Kumar. (2021). "Individually vulnerable, collectively safe: The security and privacy practices of households with older adults". *Proceedings of the ACM on Human-Computer Interaction*. 5(CSCW1): 1–24.

Napoli, D., S. N. Chaparro, S. Chiasson, and E. Stobert. (2020). "Something Doesn't Feel Right: Using Thermal Warnings to Improve User Security Awareness".

Ng, W. (2012). "Can we teach digital natives digital literacy?" *Computers & education*. 59(3): 1065–1078.

Nicholson, J., L. Coventry, and P. Briggs. (2019). "" If It's Important It Will Be A Headline" Cybersecurity Information Seeking in Older Adults". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

Nicholson, J., V. Vlachokyriakos, L. Coventry, P. Briggs, and P. Olivier. (2018). "Simple nudges for better password creation". In: *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32*. 1–12.

Nielsen, J. and J. Alertbox. (2004). "User education is not the answer to security problems". *Alertbox, October*.

Norberg, P. A., D. R. Horne, and D. A. Horne. (2007). "The privacy paradox: Personal information disclosure intentions versus behaviors". *Journal of consumer affairs*. 41(1): 100–126.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.

Ohlhausen, M. K. (2014). "Privacy challenges and opportunities: The role of the Federal Trade Commission". *Journal of Public Policy & Marketing*. 33(1): 4–9.

Ohyama, T. and A. Kanaoka. (2015). "Password strength meters using social influence". In: *Proceedings of the Symposium on Usable Privacy and Security (SOUPS'15). Usenix, Berkely, CA*.

Olmstead, K. and A. Smith. (2017). "Americans and Cybersecurity". *Tech. rep.* Pew Research Center. URL: http://www.pewinternet.org/2017/01/26/americans-and-cybersecurity/.

Owens, K., O. Anise, A. Krauss, and B. Ur. (2021). "User Perceptions of the Usability and Security of Smartphones as FIDO2 Roaming Authenticators". In: *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 57–76.

Park, C. Y., C. Faklaris, S. Zhao, A. Sciuto, L. Dabbish, and J. Hong. (2018). "Share and share alike? An exploration of secure behaviors in romantic relationships". In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 83–102.

Parsons, K., D. Calic, M. Pattinson, M. Butavicius, A. McCormac, and T. Zwaans. (2017). "The human aspects of information security questionnaire (HAIS-Q): two further validation studies". *Computers & Security*. 66: 40–51.

Pearman, S., S. A. Zhang, L. Bauer, N. Christin, and L. F. Cranor. (2019). "Why people (don't) use password managers effectively". In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 319–338.

Petelka, J., Y. Zou, and F. Schaub. (2019). "Put your warning where your link is: Improving and evaluating email phishing warnings". In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–15.

Rader, E., R. Wash, and B. Brooks. (2012). "Stories as informal lessons about security". In: *Proceedings of the Eighth Symposium on Usable Privacy and Security*. 1–17.

Redmiles, E. (2018). "Net benefits: Digital inequities in social capital, privacy preservation, and digital parenting practices of US social media users". In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 12. No. 1.

Redmiles, E. M., S. Kross, and M. L. Mazurek. (2016a). "How i learned to be secure: a census-representative survey of security advice sources and behavior". In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. 666–677.

Redmiles, E. M., S. Kross, and M. L. Mazurek. (2017). "Where is the digital divide? a survey of security, privacy, and socioeconomics". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 931–936.

Redmiles, E. M., A. R. Malone, and M. L. Mazurek. (2016b). "I think they're trying to tell me something: Advice sources and selection for digital security". In: *2016 IEEE Symposium on Security and Privacy (SP)*. IEEE. 272–288.

Redmiles, E. M., M. L. Mazurek, and J. P. Dickerson. (2018). "Dancing pigs or externalities? Measuring the rationality of security decisions". In: *Proceedings of the 2018 ACM Conference on Economics and Computation*. 215–232.

Reeder, R. W., L. Bauer, L. F. Cranor, M. K. Reiter, and K. Vaniea. (2011). "More than skin deep: measuring effects of the underlying model on access-control system usability". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2065–2074.

Reeder, R. W., L. Bauer, L. F. Cranor, M. K. Reiter, K. Bacon, K. How, and H. Strong. (2008). "Expandable grids for visualizing and authoring computer security policies". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1473–1482.

Reisig, M. D., D. P. Mears, S. E. Wolfe, and K. Holtfreter. (2015). "Financial Exploitation of the Elderly in a Consumer Context".

Rezgui, Y. and A. Marks. (2008). "Information security awareness in higher education: An exploratory study". *Computers & security*. 27(7-8): 241–253.

Roca, J. C., J. J. García, and J. J. De La Vega. (2009). "The importance of perceived trust, security and privacy in online trading systems". *Information Management & Computer Security*.

Roepke, R. and U. Schroeder. (2019). "The Problem with Teaching Defence against the Dark Arts: A Review of Game-based Learning Applications and Serious Games for Cyber Security Education." *CSEDU (2)*: 58–66.

Rogers, E. M. (1962). *Diffusion of innovations*. Free Press of Glencoe. URL: http://books.google.com/books?id=zw0-AAAAIAAJ.

Rogers, E. M. (2002). "Diffusion of preventive innovations". *Addictive Behaviors*. 27: 989–993.

Ruoti, S., J. Andersen, T. Hendershot, D. Zappala, and K. Seamons. (2016). "Private Webmail 2.0: Simple and easy-to-use secure email". In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 461–472.

Ruoti, S., J. Andersen, D. Zappala, and K. Seamons. (2015). "Why Johnny still, still can't encrypt: Evaluating the usability of a modern PGP client". *arXiv preprint arXiv:1510.08555*.

Ruoti, S. and K. Seamons. (2019). "Johnny's journey toward usable secure email". *IEEE Security & Privacy*. 17(6): 72–76.

Sadeh, N., J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao. (2009). "Understanding and capturing people's privacy policies in a mobile social networking application". *Personal and ubiquitous computing*. 13(6): 401–412.

Sasse, M. (2003). "Computer security: Anatomy of a Usability Disaster, and a Plan for Recovery". In: *Proc. CHI Workshop on HCI and Security Systems*. Citeseer. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.9019&rep=rep1&type=pdf.

Schaub, F., R. Balebako, A. L. Durity, and L. F. Cranor. (2015). "A design space for effective privacy notices". In: *Eleventh symposium on usable privacy and security (SOUPS 2015)*. 1–17.

Schechter, S., S. Egelman, and R. W. Reeder. (2009). "It's not what you know, but who you know: a social approach to last-resort authentication". In: *Proceedings of the sigchi conference on human factors in computing systems*. 1983–1992.

Schneier, B. (2015). *Secrets and lies: digital security in a networked world*. John Wiley & Sons.

Scrimgeour, J.-M. and J. Ophoff. (2019). "Lessons learned from an organizational information security awareness campaign". In: *IFIP World Conference on Information Security Education*. Springer. 129–142.

Shan, S., E. Wenger, J. Zhang, H. Li, H. Zheng, and B. Y. Zhao. (2020). "Fawkes: Protecting Privacy against Unauthorized Deep Learning Models". *In Proc. of the 29th USENIX Security Symposium*.

Shay, R., L. Bauer, N. Christin, L. F. Cranor, A. Forget, S. Komanduri, M. L. Mazurek, W. Melicher, S. M. Segreti, and B. Ur. (2015). "A spoonful of sugar? The impact of guidance and feedback on password-creation behavior". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2903–2912.

Sheng, S., L. Broderick, C. A. Koranda, and J. J. Hyland. (2006). "Why johnny still can't encrypt: evaluating the usability of email encryption software". In: *Symposium On Usable Privacy and Security*. ACM. 3–4.

Sheng, S., B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge. (2007). "Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish". In: *Proceedings of the 3rd symposium on Usable privacy and security*. 88–99.

Shneiderman, B. and P. Maes. (1997). "Direct manipulation vs. interface agents". *interactions*. 4(6): 42–61.

Siponen, M. T. (2000). "A conceptual foundation for organizational information security awareness". *Information management & computer security*.

Smith, H. J., S. J. Milberg, and S. J. Burke. (1996). "Information privacy: Measuring individuals' concerns about organizational practices". *MIS quarterly*: 167–196.

Smith, Z. M. and E. Lostri. (2020). *The hidden costs of cybercrime*. McAfee.

Solove, D. J. (2007). "I've got nothing to hide and other misunderstandings of privacy". *San Diego L. Rev.* 44: 745.

Song, Y., C. Faklaris, Z. Cai, J. I. Hong, and L. Dabbish. (2019). "Normal and easy: Account sharing practices in the workplace". *Proceedings of the ACM on Human-Computer Interaction*. 3(CSCW): 1–25.

Spiekermann, S. (2007). "Perceived control: Scales for privacy in ubiquitous computing". In: *Digital Privacy*. Auerbach Publications. 289–304.

Stanton, J., P. Mastrangelo, K. Stam, and J. Jolton. (2004). "Behavioral Information Security: Two End User Survey Studies of Motivation and Security Practices." *AMCIS*. (August): 2–8. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.92.2938&rep=rep1&type=pdf.

Stobert, E. and R. Biddle. (2014). "A password manager that doesn't remember passwords". In: *Proceedings of the 2014 New Security Paradigms Workshop*. 39–52.

Stobert, E., T. Safaie, H. Molyneaux, M. Mannan, and A. Youssef. (2020). "ByPass: Reconsidering the usability of password managers". In: *International Conference on Security and Privacy in Communication Systems*. Springer. 446–466.

Story, P., D. Smullen, A. Acquisti, L. F. Cranor, N. Sadeh, and F. Schaub. (2020). "From intent to action: Nudging users towards secure mobile payments". In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 379–415.

Strand, K. L. (2018). "Influencing factors and effectiveness of a security awareness campaign". *MA thesis*. NTNU.

Stylios, I. C., O. Thanou, I. Androulidakis, and E. Zaitseva. (2016). "A review of continuous authentication using behavioral biometrics". In: *Proceedings of the SouthEast European Design Automation, Computer Engineering, Computer Networks and Social Media Conference*. 72–79.

Sultana, M., P. P. Paul, and M. Gavrilova. (2014). "A concept of social behavioral biometrics: motivation, current developments, and future trends". In: *2014 International Conference on Cyberworlds*. IEEE. 271–278.

Sunshine, J., S. Egelman, H. Almuhimedi, N. Atri, and L. F. Cranor. (2009). "Crying wolf: An empirical study of ssl warning effectiveness." In: *USENIX security symposium*. Montreal, Canada. 399–416.

Thaler, R. H. and C. R. Sunstein. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Thorpe, J., B. MacRae, and A. Salehi-Abari. (2013). "Usability and security evaluation of GeoPass: a geographic location-password scheme". In: *Proceedings of the Ninth symposium on usable privacy and security*. 1–14.

Tiefenau, C., M. Häring, K. Krombholz, and E. Von Zezschwitz. (2020). "Security, availability, and multiple information sources: Exploring update behavior of system administrators". In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 239–258.

Tiefenau, C., E. von Zezschwitz, M. Häring, K. Krombholz, and M. Smith. (2019). "A usability evaluation of Let's Encrypt and Certbot: usable security done right". In: *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. 1971–1988.

Ur, B., F. Alfieri, M. Aung, L. Bauer, N. Christin, J. Colnago, L. F. Cranor, H. Dixon, P. Emami Naeini, H. Habib, *et al.* (2017). "Design and evaluation of a data-driven password meter". In: *Proceedings of the 2017 chi conference on human factors in computing systems*. 3775–3786.

Vance, A., J. L. Jenkins, B. B. Anderson, D. K. Bjornn, and C. B. Kirwan. (2018). "Tuning out security warnings: A longitudinal examination of habituation through fMRI, eye tracking, and field experiments". *MIS Quarterly*. 42(2): 355–380.

Vance, A., B. Kirwan, D. Bjornn, J. Jenkins, and B. B. Anderson. (2017). "What do we really know about how habituation to warnings occurs over time? A longitudinal fMRI study of habituation and polymorphic warnings". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2215–2227.

Vaniea, K., L. Bauer, L. F. Cranor, and M. K. Reiter. (2012). "Out of sight, out of mind: Effects of displaying access-control information near the item it controls". In: *2012 Tenth Annual International Conference on Privacy, Security and Trust*. IEEE. 128–136.

Vaniea, K. and Y. Rashidi. (2016). "Tales of software updates: The process of updating software". In: *Proceedings of the 2016 chi conference on human factors in computing systems*. 3215–3226.

Vaniea, K. E., E. Rader, and R. Wash. (2014). "Betrayed by updates: how negative experiences affect future security". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2671–2674.

Vannini, S., R. Gomez, and B. C. Newell. (2020). ""Mind the five": Guidelines for data privacy and security in humanitarian work with undocumented migrants and other vulnerable populations". *Journal of the Association for Information Science and Technology*. 71(8): 927–938.

Venkatesh, V. and H. Bala. (2008). "Technology acceptance model 3 and a research agenda on interventions". *Decision sciences*. 39(2): 273–315.

Venkatesh, V. and F. D. Davis. (2000). "A theoretical extension of the technology acceptance model: Four longitudinal field studies". *Management science*. 46(2): 186–204.

Venkatesh, V., M. G. Morris, G. B. Davis, and F. D. Davis. (2003). "User acceptance of information technology: Toward a unified view". *MIS quarterly*: 425–478.

Verizon. (2020). "Data Breach Investigations Report". *Tech. rep.*

Verizon. (2022). "Data Breach Investigations Report". *Tech. rep.*

Wang, Y.-C., R. Kraut, and J. M. Levine. (2012). "To stay or leave? The relationship of emotional and informational support to commitment in online health support groups". In: *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 833–842.

Wang, Q., H. Jin, and N. Li. (2009). "Usable access control in collaborative environments: Authorization based on people-tagging". In: *European Symposium on Research in Computer Security*. Springer. 268–284.

Wang, S., C. Faklaris, J. Lin, L. Dabbish, and J. I. Hong. (2022). "'It's Problematic but I'm not Concerned': University Perspectives on Account Sharing". *Proceedings of the ACM on Human-Computer Interaction*. 6(CSCW1): 1–27.

Wang, Y. (2018). "Inclusive security and privacy". *IEEE Security & Privacy*. 16(4): 82–87.

Wang, Y., P. G. Leon, A. Acquisti, L. F. Cranor, A. Forget, and N. Sadeh. (2014). "A field trial of privacy nudges for facebook". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2367–2376.

Wash, R. (2010). "Folk models of home computer security". In: *Proc. SOUPS '10*. New York, New York, USA: ACM Press. 1. DOI: 10.1145/1837110.1837125.

Wash, R., E. Rader, K. Vaniea, and M. Rizor. (2014). "Out of the loop: How automated software updates cause unintended security consequences". In: *10th Symposium On Usable Privacy and Security (SOUPS) 2014)*. 89–104.

WebAIM. (2017). "Screen Reader User Survey #7 Results".

Whitten, A. and J. Tygar. (1999). "Why Johnny can't encrypt: A usability evaluation of PGP 5.0". In: *Proc. SSYM'99*. 14–28. URL: http://www.usenix.org/events/sec99/full_papers/whitten/whitten.ps.

Wickins, J. (2007). "The ethics of biometrics: the risk of social exclusion from the widespread use of electronic identification". *Science and Engineering Ethics*. 13(1): 45–54.

Wikipedia. (2021). "2017 Equifax data breach". URL: https://en.wikipedia.org/wiki/2017_Equifax_data_breach.

Wilson, G., H. Maxwell, and M. Just. (2017). "Everything's Cool: Extending Security Warnings with Thermal Feedback". In: *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*. 2232–2239.

Wobbrock, J. O. (2009). "Tapsongs: tapping rhythm-based passwords on a single binary sensor". In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. 93–96.

Wogalter, M. S. (2006a). "Behavioral compliance: Theory, methodology, and results". In: *Handbook of warnings*. CRC Press. 335–354.

Wogalter, M. S. (2006b). “Communication-human information processing (C-HIP) model”. *Handbook of warnings*: 51–61.

Woo, S., E. Kaiser, R. Artstein, and J. Mirkovic. (2016). “Life-experience passwords (leps)”. In: *Proceedings of the 32nd Annual Conference on Computer Security Applications*. 113–126.

Wu, Y., W. K. Edwards, and S. Das. (2022a). ““A Reasonable Thing to Ask For”: Towards a Unified Voice in Privacy Collective Action”. In: *CHI Conference on Human Factors in Computing Systems*. 1–17.

Wu, Y., W. K. Edwards, and S. Das. (2022b). “SoK: Social Cybersecurity”. In: *IEEE Symposium on Security and Privacy (Oakland)(2022). https://sauvikdas. com/uploads/paper/pdf/36/file. pdf*.

Xiao, S., J. Witschey, and E. Murphy-Hill. (2014). “Social influences on secure development tool adoption: why security tools spread”. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1095–1106.

Xu, H., S. Gupta, M. B. Rosson, and J. M. Carroll. (2012). “Measuring mobile users’ concerns for information privacy”.

Yao, Y., D. Lo Re, and Y. Wang. (2017). “Folk models of online behavioral advertising”. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1957–1969.

Yıldırım, M. and I. Mackie. (2019). “Encouraging users to improve password security and memorability”. *International Journal of Information Security*. 18(6): 741–759.

Zhang, Z., Z. Zhang, H. Yuan, N. M. Barbosa, S. Das, and Y. Wang. (2021). “WebAlly: Making Visual Task-based CAPTCHAs Transferable for People with Visual Impairments”. In: *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 281–298.

Zhao, Y., J. Ye, and T. Henderson. (2014). “Privacy-aware location privacy preference recommendations”. In: *Proceedings of the 11th international conference on mobile and ubiquitous systems: Computing, networking and services*. 120–129.

Zuboff, S. (2015). “Big other: surveillance capitalism and the prospects of an information civilization”. *Journal of information technology*. 30(1): 75–89.

Zurko, M. E. and R. T. Simon. (1996). "User-centered security". In: *Proceedings of the 1996 workshop on New security paradigms*. 27–33.